

Bayer, Sonja

Mathematikunterricht im Vergleich zwischen den Schularten

Münster ; New York : Waxmann 2020, 324 S. - (Empirische Erziehungswissenschaft; 76) - (Dissertation, Universität Frankfurt, 2019)



Quellenangabe/ Reference:

Bayer, Sonja: Mathematikunterricht im Vergleich zwischen den Schularten. Münster ; New York : Waxmann 2020, 324 S. - (Empirische Erziehungswissenschaft; 76) - (Dissertation, Universität Frankfurt, 2019) - URN: urn:nbn:de:0111-pedocs-210703 - DOI: 10.25656/01:21070

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-210703>

<https://doi.org/10.25656/01:21070>

in Kooperation mit / in cooperation with:



WAXMANN
www.waxmann.com

<http://www.waxmann.com>

Nutzungsbedingungen

Dieses Dokument steht unter folgender Creative Commons-Lizenz:
<http://creativecommons.org/licenses/by-nc-sa/4.0/deed.de> - Sie dürfen das Werk bzw. den Inhalt unter folgenden Bedingungen vervielfältigen, verbreiten und öffentlich zugänglich machen sowie Abwandlungen und Bearbeitungen des Werkes bzw. Inhaltes anfertigen: Sie müssen den Namen des Autors/Rechteinhabers in der von ihm festgelegten Weise nennen. Dieses Werk bzw. der Inhalt darf nicht für kommerzielle Zwecke verwendet werden. Die neu entstandenen Werke bzw. Inhalte dürfen nur unter Verwendung von Lizenzbedingungen weitergegeben werden, die mit denen dieses Lizenzvertrages identisch oder vergleichbar sind.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

This document is published under following Creative Commons-Licence:
<http://creativecommons.org/licenses/by-nc-sa/4.0/deed.en> - You may copy, distribute and transmit, adapt or exhibit the work in the public and alter, transform or change this work as long as you attribute the work in the manner specified by the author or licensor. You are not allowed to make commercial use of the work. If you alter, transform, or change this work in any way, you may distribute the resulting work only under this or a comparable license.

By using this particular document, you accept the above-stated conditions of use.



Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Mitglied der


Leibniz-Gemeinschaft

Sonja Bayer

Mathematikunterricht im Vergleich zwischen den Schularten

Empirische Erziehungswissenschaft

herausgegeben von

Rolf Becker, Sigrid Blömeke, Wilfried Bos,
Hartmut Ditton, Cornelia Gräsel, Eckhard Klieme,
Kai Maaz, Thomas Rauschenbach, Hans-Günther Roßbach,
Knut Schwippert, Ludwig Stecher, Christian Tarnai,
Rudolf Tippelt, Rainer Watermann, Horst Weishaupt

Band 76

Sonja Bayer

Mathematikunterricht im Vergleich zwischen den Schularten



Waxmann 2020
Münster • New York

Der Fachbereich Erziehungswissenschaften der Johann Wolfgang Goethe Universität zu Frankfurt am Main hat diese Arbeit unter dem Titel „Alle alles ganz lehren – aber wie? Mathematikunterricht vergleichend zwischen den Schular-ten“ als Inaugural-Dissertation zur Erlangung der Doktorwürde im Fach Erzie- hungswissenschaften im Jahr 2019 angenommen.

Die Publikation wurde gefördert durch den Open-Access-Publikationsfonds für Monografien der Leibniz-Gemeinschaft.

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.dnb.de> abrufbar.

D.30

Empirische Erziehungswissenschaft, Bd. 76

Print-ISBN 978-3-8309-4247-4

E-Book-ISBN 978-3-8309-9247-9

DOI: 10.31244/9783830992479

© Waxmann Verlag GmbH, 2020

www.waxmann.com

info@waxmann.com

Umschlaggestaltung: Pleßmann Kommunikationsdesign, Ascheberg

Satz: Finn Volpert, DIPF, Frankfurt am Main

Dieses E-Book ist unter der Lizenz CC BY-NC-SA open access verfügbar.



Danksagung

Die vorliegende Dissertation ist im Rahmen meiner Tätigkeit als wissenschaftliche Mitarbeiterin am DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation entstanden. Auf diesem Weg haben mich verschiedene Personen begleitet und unterstützt. Ich möchte mich bei all diesen Personen herzlich bedanken.

An erster Stelle möchte ich mich bei meinem Betreuer und Gutachter, Professor Dr. Eckhard Klieme, bedanken: für die Entwicklungsmöglichkeiten, die er mir gegeben hat, für die Freiheiten bei der Themenfindung sowie für die reflektierenden und lehrreichen Gespräche über theoretische und methodische Fragen, die mir wertvolle Denkanstöße für diese Arbeit gegeben haben.

Ein weiterer Dank gilt meiner Betreuerin, Dr. Svenja Vieluf. Sie war bei inhaltlichen und methodischen Fragen der Arbeit eine wichtige Ansprechpartnerin. Professor Dr. Tobias Feldhoff danke ich für die spannenden Gespräche zu Schulkulturen und Schulentwicklung sowie für das Zweitgutachten zu meiner Dissertation.

Vor und während meiner Promotionszeit war ich in das PISA-Team eingebunden. Insbesondere bei meinen PISA-Kollegen am DIPF möchte ich mich herzlich für die tolle Teamarbeit und die netten gemeinsamen Pausen bedanken. Mein besonderer Dank gilt Dr. Nina Rothen, Dr. Nina Jude und Ingrid Mader. Sie haben mir wichtige Impulse für die Promotion und darüber hinaus gegeben. Die Kolleginnen und Kollegen im internationalen PISA-Kontext haben meinen Blick auf Bildungssysteme und Kulturen erweitert und mir neue Sichtweisen erschlossen, wovon diese Arbeit merklich profitiert hat. Vielen Dank für die spannenden Gespräche und den tollen Austausch.

Über den Projektkontext hinaus bietet das DIPF zahlreiche Austauschmöglichkeiten in informellen Runden, Kolloquien und Netzwerken. Allen hier beteiligten Kolleginnen und Kollegen möchte ich herzlich für die hilfreichen Rückmeldungen und den Austausch danken, ganz besonders Dr. Julia Dohrmann, Dr. Brigitte Steinert, Julia Käfer und Prof. Dr. Markus Sauerwein. Vielen Dank auch an Finn Volpert für seine ausdauernde Hilfe bei der formalen Erstellung der Dissertationsschrift sowie den Kolleginnen und Kollegen im Zentrum für

internationale Vergleichsstudien für die spannenden Daten, die im Verbundkontext entstanden und Grundlage meiner Analysen sind.

Abschließend danksagen möchte ich noch meiner Familie, meiner Tochter und meinen Freunden. Insbesondere Michaela Opp, Tilo Trotzke, Esther Diering und Tim Skrzypek haben mich unterstützt, aber auch alle anderen standen mir beiseite. Ohne eure Unterstützung wäre diese Arbeit nie entstanden! Vielen Dank!

Inhalt

1	Einleitung.....	11
2	Das Bildungssystem der Sekundarstufe I	16
2.1	Ein historischer Abriss	17
2.1.1	Das dreigliedrige Bildungssystem entwickelt sich	18
2.1.2	Das dreigliedrige Bildungssystem im Wandel	23
2.2	Das Bildungssystem der Sekundarstufe I heute.....	26
2.2.1	Hauptschule.....	28
2.2.2	Realschule	30
2.2.3	Gymnasium	33
2.2.4	Bildungsgangübergreifende Schulen.....	35
2.2.5	Zusammenfassende Gegenüberstellung der Schularten der Sekundarstufe I.....	42
3	Schulartspezifische Ausgangsbedingungen für den Unterricht	44
3.1	Unterricht und Unterrichtsprozesse	48
3.1.1	Unterrichtsmethoden	50
3.1.2	Formatives Assessment	53
3.1.3	Generische Faktoren des Unterrichts.....	54
3.2	Lehrende.....	57
3.2.1	Professionelle Kompetenz von Lehrenden	59
3.2.2	Erwerb professioneller Kompetenz	62
3.3	Gegenstand	66
3.3.1	Lehrpläne und Curricula – eine Begriffsbestimmung.....	67
3.3.2	Lehrpläne und Curricula der Schularten.....	69
3.3.3	Schulische Lerngelegenheiten	71

3.4	Schülergruppe.....	75
3.4.1	Exkurs: Soziale Herkunft und schulische Normen, Werte und Einstellungen?	77
3.4.2	Normative Peer-Effekte.....	81
3.4.3	Instrumentelle Peer-Effekte.....	87
3.4.4	Big-Fish-Little-Pond-Effekt	90
3.4.5	Erwartungen der Lehrenden	92
3.5	Lehren und Lernen vergleichend zwischen den Schularten – der Forschungsstand.....	95
3.5.1	Lernende: Leistung, psychosoziale Merkmale und Differenzielle Lern- und Entwicklungsmilieus	97
3.5.2	Lehrende und ihre Überzeugungen zum Lehren und Lernen	102
3.5.3	Gegenstand: Unterrichtszeit und Unterrichtsinhalte.....	106
3.5.4	Unterrichtsprozesse	108
4	Herausforderungen des Vergleichs von Schularten	115
4.1	Formen von Bias	119
4.2	Antworttendenzen.....	121
4.3	Implikationen für den Vergleich von Schularten.....	123
5	Herleitung der empirischen Fragestellungen	125
6	Methode.....	133
6.1	Stichprobe.....	134
6.2	Instrumente	135
6.2.1	Skalen, Indizes und Leistungsdaten der Lernenden.....	136
6.2.2	Skalen der Lehrenden	138
6.3	Analyseebene und Gewichtung	139
6.4	Vergleichbarkeit der Fragebogeninstrumente.....	141
6.4.1	Invarianzanalyse mittels mehrgruppenkonfirmatorischer Faktorenanalyse.....	141
6.4.2	Faktorstrukturvergleich	145

6.4.3	Berechnung von Indikatoren für Antworttendenzen	148
6.5	Varianzanalytische Modelle für die Analyse von Antworttendenzen und die Deskription der Schularten.....	150
6.6	Einfaktorielle Varianzanalyse.....	151
6.7	Kovarianzanalyse	153
6.8	Mediationsanalysen: Vermittelnde Mechanismen für Schulartunterschiede in den Unterrichtsprozessen	154
6.9	Moderationsanalysen: Schulartunterschiede in den Zusammenhängen zwischen Unterrichtsprozessen und Leistungsentwicklung	158
7	Ergebnisse.....	162
7.1	Vergleichbarkeit von Daten zum Lehren und Lernen aus Fragebogenstudien	163
7.1.1	Äquivalenz der Messinstrumente	164
7.1.2	Antworttendenzen und deren Effekte auf Mittelwerte	178
7.1.3	Zusammenfassung der Ergebnisse zur Vergleichbarkeit und Schlussfolgerungen für die weiteren Analysen	185
7.2	Deskription der Schularten	187
7.2.1	Zusammensetzung der Schülergruppe.....	188
7.2.2	Kognitives Lernumfeld.....	189
7.2.3	Sozioökonomisches Lernumfeld	194
7.2.4	Normatives Lernumfeld.....	196
7.2.5	Überzeugungen der Lehrenden zu den Zielen im Mathematikunterricht	199
7.2.6	Inhalte des Mathematikunterrichts	202
7.2.7	Unterrichtsprozesse im Fach Mathematik	205
7.2.8	Zusammenfassende Interpretation	213

7.3	Schulartunterschiede in den Unterrichtsprozessen mediiert über die Überzeugungen der Lehrenden, die Unterrichtsinhalte und die Schülergruppe	216
7.3.1	Potential zur kognitiven Aktivierung durch Variation der Aufgaben	218
7.3.2	Schülerorientierte Unterrichtsmethoden.....	219
7.3.3	Zusammenfassende Interpretation	223
7.4	Schulartunterschiede in den Zusammenhängen zwischen Unterrichtsprozessen und Leistungsentwicklung: Moderationseffekte der Schulart	225
7.4.1	Zusammenhänge von Unterrichtsprozessmerkmalen und Leistungsentwicklung gemessen mit dem BiSta-Test	227
7.4.2	Zusammenhänge von Unterrichtsprozessmerkmalen und Leistungsentwicklung gemessen mit dem PISA-Test	228
7.4.3	Zusammenfassende Interpretation	230
8	Diskussion	232
8.1	Bedeutung der Schulartunterschiede	234
8.2	Vergleichbarkeit von Fragebogendaten: Implikationen für die Schul- und Unterrichtsforschung.....	240
8.3	Grenzen des empirischen Ansatzes der Arbeit und inhaltliche Einschränkungen	244
8.4	Schulstruktur oder Einzelschule – was zählt?.....	249
8.5	Schularten im Kontext gesellschaftlicher, wirtschaftlicher und individueller Anforderungen und Entwicklungen	251
9	Literaturverzeichnis	258
10	Abbildungsverzeichnis	289
11	Tabellenverzeichnis	291

Der Anhang ist abrufbar unter www.waxmann.com/buch4247.

1 Einleitung

„Alle alles ganz zu lehren“ sei das Ziel jeglicher Bildungsbemühungen, postuliert Johann Amos Comenius in seiner „Didactica magna“ aus dem Jahr 1657. Sein Konzept der Bildungsorganisation sieht dabei die gemeinsame Beschulung aller vor – unabhängig von Geschlecht, Stand oder Begabung. Erst mit der Vorbereitung auf die wissenschaftliche Laufbahn in der Lateinschule sollten sich nach Comenius’ Entwurf die Bildungswege trennen (Kuhlmann 2013; Raithel et al. 2007). Eine einheitliche Bildung schulpflichtiger Kinder ist in Deutschland jedoch bis heute nicht realisiert. Stattdessen trennen sich nach der gemeinsamen Grundschule die Bildungswege. Somit wird dem Anrecht auf Bildung aller in unterschiedlichen Schularten mit je eigenen Bildungszielen und Schwerpunkten Rechnung getragen (Autorengruppe Bildungsberichterstattung 2016). Hier lassen sich fünf übergreifende Schularten der Sekundarstufe I abgrenzen: Hauptschulen, Realschulen, Gymnasien, Schulen mit mehreren Bildungsgängen und integrierte Gesamtschulen.

Welche Schulart die Lernenden besuchen, entscheidet – entgegen der Forderung Comenius’, allen die gleiche Allgemeinbildung zu ermöglichen, – maßgeblich über den weiteren Bildungsverlauf (Ditton und Krüskens 2010). Dies bezieht sich nicht nur auf Unterschiede zwischen den Schularten in den angebotenen Fächern und auf die Wahrscheinlichkeit, einen bestimmten Abschluss zu erzielen, sondern auch auf die kognitiven und affektiven Entwicklungsmöglichkeiten der Lernenden: Vergleichende Untersuchungen legen offen, dass sich die Leistungsentwicklung der Lernenden in allgemeinen Kompetenzbereichen wie Mathematik oder Lesen zwischen den Schularten unterscheiden und zwar auch dann, wenn die individuellen Eingangsleistungen berücksichtigt werden (Becker et al. 2012; Guill et al. 2017; Pfof et al. 2010). Dies bedeutet, dass Lernende mit der gleichen Eingangsleistung abhängig von der Schulart Unterschiede im Kompetenzerwerb zeigen. Die größten Unterschiede in der Leistungsentwicklung zeigen sich bei den mathematischen Fähigkeiten, welche im Vergleich zum Lesevermögen besonders stark durch das schulische Lernen geprägt sind. Auch für die Entwicklung affektiver Merkmale wie Motivation, Selbstkonzept oder Interesse finden einzelne Studien Unterschiede zwischen den Schularten (bspw. Becker und Neumann 2016; Wagner et al. 2008). Die jeweiligen Bedingungen von Schule und Unterricht scheinen also unterschiedlich förderlich für die kognitive wie auch die nichtkognitive Entwicklung der Lernenden zu sein.

Es stellt sich die Frage, was zu den unterschiedlichen Entwicklungsmöglichkeiten der Lernenden innerhalb der Schularten führt. Aus der Forschung zu Kompositionseffekten, welche den Einfluss der Schülergruppe auf die Leistungsentwicklung des einzelnen Lernenden untersucht, kann gefolgert werden, dass die spezifische Zusammensetzung der Schülergruppe an den Schularten eine Ursache sein kann (Dumont et al. 2013). In Abgrenzung zur Komposition werden in der Literatur Merkmale der Institution als Ursache für unterschiedliche Entwicklungsmöglichkeiten an den verschiedenen Schularten genannt. Hier gelten der Lehrplan, die Lehrerbildung sowie die Unterrichtsziele und -methoden als Einflussfaktoren für die Entwicklungsmöglichkeiten der Lernenden (bspw. Baumert et al. 2006b; Baumert et al. 2011; Becker et al. 2012; Scharenberg 2012). Jedoch mangelt es bislang an theoretischen Überlegungen und Studien, welche die zugrunde liegenden Mechanismen erklären und analysieren.

Ziel der hier vorliegenden Arbeit ist es daher, zum einen den Unterricht vergleichend zwischen den Schularten zu betrachten und zum anderen für Unterschiede in den Unterrichtsprozessen potentielle Vermittlungsmechanismen zu untersuchen. Um relevante Merkmale herzuleiten, verknüpft die hier vorliegende Arbeit die Systemebene mit theoretischen Modellen und Befunden aus der empirischen Unterrichtsforschung. Konkret werden Merkmale der Unterrichtsinhalte, der Lehrerüberzeugungen sowie der Schülergruppe als Vermittlungsmechanismen untersucht. Zudem soll ergründet werden, welche Bedeutung die Unterschiede in den Unterrichtsprozessen für die kognitive Entwicklung der Lernenden haben. Für die empirischen Analysen werden Daten der nationalen PISA-2012-Studie (Programme for International Student Assessment) herangezogen, welche den Schwerpunkt auf den Mathematikunterricht und die Mathematikleistung von Lernenden legt. Die Fragestellungen und Analysen beziehen sich daher konkret auf den Mathematikunterricht von Neuntklässlern. Da die Beschreibung des Unterrichts auf Fragebogendaten beruht, wird zudem untersucht, ob die eingesetzten Fragebogeninstrumente invariant zwischen den Schularten sind und inwieweit die Daten durch Antworttendenzen verzerrt sind. Somit wird vorab geprüft, ob die verwendeten Daten der PISA-2012-Studie geeignet sind, um die geplanten Schulartvergleiche anzustellen.

Die für die Analysen verwendeten Messinstrumente auf Invarianz zwischen den Schularten zu prüfen sowie Antworttendenzen zu berücksichtigen, scheint beim Vergleich von Schularten bislang einmalig zu sein. Eine weitere Besonderheit

dieser Arbeit liegt in den verwendeten Leistungstests. Neben den Leistungsdaten aus dem PISA-Test werden in parallelen Analysen auch die Leistungsdaten der curriculumsbezogenen Tests zur Überprüfung der Bildungsstandards in die Analysen einbezogen. An dieser Stelle sei vorweggenommen, dass sich die Zusammenhänge der untersuchten Merkmale mit der Mathematikleistung nicht nur teilweise zwischen den Schularten sondern auch teilweise zwischen den beiden hier verwendeten Leistungstests unterscheiden.

Die Erkenntnisse aus dieser Arbeit sind in mehrfacher Sicht bedeutend für die empirische Bildungsforschung. Zum einen beleuchten die Ergebnisse zur Invarianz der Messinstrumente und zum Einfluss von Antworttendenzen Einschränkungen bei der Analyse und Ergebnisinterpretation von Fragebogenstudien in der Schul- und Unterrichtsforschung. Zum anderen erweitern der Vergleich der Schularten sowie die Analyse von Vermittlungsmechanismen für Schulartunterschiede in den Unterrichtsprozessen das Verständnis des Mathematikunterrichts und dessen Kontextgebundenheit. Damit könnten die Erkenntnisse aus dieser Arbeit auch fruchtbar für die Interpretation von Ergebnissen anderer Studien sein – auch für jene Studien, welche bei der Analyse keine Unterschiede zwischen Schularten einbeziehen. Die parallelen Analysen mit zwei unterschiedlichen Leistungstests regen zudem an, die getesteten Inhalte von Leistungstests bei der Interpretation von Ergebnissen einzubeziehen. Auch für die Frage nach der Schulstruktur in Deutschland soll diese Arbeit einen Beitrag leisten. Zum einen wird aufgezeigt, dass mit den Schularten nicht nur strukturelle sondern auch inhaltliche Konsequenzen verbunden sind. Die vorliegende Arbeit möchte damit dazu beitragen, die Schulstrukturfrage wieder aktiv auf nationaler Ebene anzugehen und auch die Ziele und Inhalte, die mit den Schularten verbunden sind, kritisch hinsichtlich der gesellschaftlichen und wirtschaftlichen Erfordernisse und Entwicklungen zu diskutieren. Zum anderen bietet die hier vorliegende Arbeit für die wissenschaftliche Begleitforschung von Schulstrukturreformen in einzelnen Bundesländern Anregungen für relevante Merkmale, welche in die Untersuchungen der Begleitforschung einbezogen werden könnten.

Die hier vorliegende Arbeit bezieht theoretische Überlegungen sowie empirische Erkenntnisse in die Untersuchung der Schularten ein. Konkret ist die Arbeit wie folgt aufgebaut: In Kapitel 2 wird die Entwicklung der Schularten der Sekundarstufe I nachgezeichnet und Unterschiede in der aktuellen Ausgestaltung zwischen den Schularten werden herausgearbeitet. Besonders in den Bil-

ungszielen und -inhalten, den Leitbildern und in den Selektionsprozessen sind Unterschiede zwischen den Schularten deutlich.

Kapitel 3 schließt mit Überlegungen an, wie sich diese systembedingten Unterschiede auf den Mathematikunterricht der Lernenden auswirken können. Im Rückgriff auf didaktische Theorien wird ein erweitertes Didaktisches Dreieck hergeleitet, welches die Vermittlungsmechanismen über die Lehrenden, den Gegenstand sowie die Schülergruppe verdeutlicht. Darauf aufbauend wird das Verständnis von Unterricht, das dieser Arbeit zugrunde liegt, dargelegt und tiefergehend begründet, wie und warum sich die Überzeugungen der Lehrenden, die Unterrichtsinhalte sowie die Gruppe der Lernenden zwischen den Schularten unterscheiden. Das Kapitel schließt mit dem Forschungsstand zu Schulartunterschieden in den Unterrichtsprozessen, den Inhalten, der Schülergruppe sowie den Überzeugungen der Lernenden und verweist auf die Forschungslücken.

Kapitel 4 reflektiert grundlegend den Vergleich der Schularten. Einleitend wird die Kritik an der vergleichenden Erziehungswissenschaft nachgezeichnet und hinsichtlich der Konsequenzen für den Vergleich von Schularten hinterfragt. Anschließend werden methodische Schwierigkeiten beim Vergleich von Schularten mit Fragebogendaten erörtert. Insbesondere werden potentielle Einschränkungen des Vergleichs durch invariante Fragebogeninstrumente und Antworttendenzen thematisiert.

Die theoretischen Überlegungen sollen schließlich empirisch betrachtet werden. Hierzu werden vier empirische Fragestellungen aufgeworfen (Kapitel 5), denen nach einer Darstellung der verwendeten Stichprobe und Methoden (Kapitel 6) in Kapitel 7 nachgegangen wird. Erstens wird geprüft, inwieweit die verwendeten Daten geeignet sind, um die Schularten vergleichend zu beschreiben und Zusammenhänge einzelner Merkmale zwischen den Schularten zu vergleichen. Zweitens werden die Schularten anhand geeigneter Merkmale hinsichtlich der Schülergruppe, der Überzeugungen von Lehrenden, der Unterrichtsinhalte und der Unterrichtsprozesse kontrastiert. Drittens wird untersucht, durch welche Merkmale der Schülergruppe, der Lehrerüberzeugungen und der Unterrichtsinhalte die Schulartunterschiede in den Unterrichtsprozessen vermittelt sind. Viertens werden die Zusammenhänge der Unterrichtsprozessmerkmale mit der Leistung der Lernenden (unter Kontrolle der Vorleistung) analysiert und ge-

prüft, inwieweit sich diese Zusammenhänge zwischen den Schularten unterscheiden.

Die Arbeit schließt mit einer Diskussion der Befunde (Kapitel 8). Zu Beginn werden die empirischen Erkenntnisse des Schulartvergleichs auf einer übergeordneten Ebene zusammengefasst. Zudem werden die Erkenntnisse zum Vergleich von Schularten anhand von Fragebogendaten verallgemeinert und Möglichkeiten besprochen, die Vergleichbarkeit von Fragebogendaten in der Schul- und Unterrichtsforschung abzusichern. Anschließend wird auf die Limitationen der vorliegenden Arbeit eingegangen. Vor diesem Hintergrund werden die empirischen Ergebnisse im Kontext der Schulstrukturdebatte diskutiert. Zum einen wird die Bedeutung der Einzelschule von der Bedeutung der Schulart abgegrenzt. Zum anderen werden die Schularten und die Schulstruktur im Kontext gesellschaftlicher, wirtschaftlicher und individueller Anforderungen reflektiert.

2 Das Bildungssystem der Sekundarstufe I

Die drei Bildungsgänge der Sekundarstufe – Hauptschulbildungsgang, Realschulbildungsgang und gymnasialer Bildungsgang – werden an verschiedenen Schularten angeboten: Hauptschulen, Realschulen, Gymnasien und bildungsgangübergreifende Schulen¹ (KMK 2014). Bei den bildungsgangübergreifenden Schulen lassen sich wiederum Schulen mit mehreren Bildungsgängen von integrierten Gesamtschulen abgrenzen (ausführlicher in Abschnitt 2.2.)

Auch Förderschulen zählen zur deutschen Bildungslandschaft. Diese nehmen ca. 4,6 % der Lernenden auf (KMK 2016a). Formal wie inhaltlich wird das Förderschulwesen jedoch von allgemeinen Schulen getrennt. Förderschulen nehmen Lernende auf, welche einen Förderbedarf aufweisen, der nicht in allgemeinen Schulen bedient werden kann (KMK 1994). Mit einem Fokus auf einem von zehn Förderschwerpunkten wird an Förderschulen die unterrichtliche Gestaltung auf die Bedürfnisse ihrer Lernenden zugeschnitten. Hierzu zählen spezielle technische und behinderungsspezifische apparative Hilfen sowie Medien, welche individuell angepasst werden können. Zudem können die Lehrpläne entsprechend individueller Fördernotwendigkeiten angepasst werden (Bundschuh und Baier 2009; KMK 1994). In ihrer inhaltlichen und formalen Variation kommt den Förderschulen häufig ein Sonderstatus zu, beispielsweise bei der Beschreibung der Bildungslandschaft (Hänsel 2005) sowie in Beschlüssen der Kultusministerkonferenz (KMK). Das Förderschulwesen kann damit als paralleles Bildungssystem betrachtet werden, das nur wenig Überschneidung mit dem Bildungssystem der allgemeinen Schulen aufweist. Auch diese Arbeit fokussiert aufgrund der Besonderheiten und hohen Variation des Förderschulwesens auf allgemeine Schulen des Bildungssystems in Deutschland.

Die starke Trennung zwischen Förderschulwesen und allgemeinen Schulen wird häufig kritisiert, besonders in der aktuell verstärkten Diskussion um Inklusion (Preuss-Lausitz 2018). Ebenso viel diskutiert und kritisiert wird die Gliederung des Bildungssystems der allgemeinen Schulen (Gehrmann 2016). Eine solch starre Struktur mit früher Weichenstellung ist im internationalen Ver-

1 Die Bezeichnungen der Schularten variieren teils zwischen den Bundesländern. Eine Übersicht findet sich im Online-Anhang 12.1 (Stand 2018).

gleich selten anzutreffen.² Ein Blick über Deutschland hinaus zeigt vielfältige Alternativformen (OECD 2017), welche in der jeweiligen Ausgestaltung das Produkt der länderspezifischen sozialen, wirtschaftlichen und politischen Aushandlungsprozesse sind (Leschinsky und Cortina 2008). Um die aktuelle Struktur des Bildungssystems in Deutschland und die damit zusammenhängenden schulischen Lerngelegenheiten innerhalb der Schularten zu verstehen, ist es daher unerlässlich, relevante Impulse und gesellschaftliche Zusammenhänge zu rekonstruieren. „In ihrer historischen Gestalt und Entwicklung folgen Schulen nicht allein den normativen Vorgaben der Bildungstheoretiker bzw. von Verfassungen und Schulgesetzen, sondern zugleich, gelegentlich zuerst, den Möglichkeiten und Zwängen, die der jeweilige Staat und eine konkrete Gesellschaft definieren“ (Diederich und Tenorth 1997, S. 45). Aus der Entstehungs- und Entwicklungsgeschichte heraus lassen sich dann die Schularten in Deutschland in ihren Zielen und Inhalten kontrastieren.

Die Beschreibung des Bildungssystems in Deutschland wird daher in der vorliegenden Arbeit über die historische Entwicklung der Schularten in Deutschland eingeleitet (Abschnitt 2.1). Der Fokus liegt auf der Entwicklungsgeschichte des allgemeinen Schulwesens in der Sekundarstufe I. Daher berücksichtigt die Beschreibung den Bereich der vorschulischen, der beruflichen und der akademischen Bildung nur dort, wo es der Zusammenhang erfordert. Aus der historischen Entwicklung heraus werden anschließend die Schularten in ihrer aktuellen Ausgestaltung beschrieben und die je eigenen Schwerpunkte in den Zielen und Inhalten herausgestellt (Abschnitt 2.2).

2.1 Ein historischer Abriss

Gesellschaftlich verankert ist das moderne Bildungssystem in Deutschland seit Mitte des 19. Jahrhunderts. Die Weitergabe von Wissen und Erziehung in mehr oder weniger formalen Settings gibt es bedeutend länger, man denke bspw. an die Sophisten im antiken Griechenland oder an Klosterschulen im Mittelalter (Protz 2004). Ein verpflichtendes Bildungswesen mit vorgegebenen Strukturen hat sich jedoch erst mit der Einführung bzw. weitgehenden Durchsetzung der Unterrichtspflicht um das Jahr 1880 voll entwickelt (Drewiek und Tenorth

2 Die Grundstruktur des Bildungssystems in Deutschland als Diagramm findet sich unter https://www.kmk.org/fileadmin/Dateien/pdf/Dokumentation/dt_2017.pdf.

2001). Die Inhalte und Ausgestaltung des Bildungssystems spiegelten die gesellschaftliche Ordnung sowie politische und soziale Entwicklungen wider. So entwickelten sich in Anlehnung an die gesellschaftlichen Strukturen die Schularten im deutschen Sprachraum auf verschiedenen Ebenen entsprechend der Bedeutung von Schule für verschiedene soziale Interessensbereiche und Schichten (Keck 2009).

Im Folgenden sei auf zentrale Personen und Ereignisse in ihrer gesellschaftlich politischen Einbettung verwiesen, welche die Entstehung und Gestaltung des Bildungssystems sowie die inhaltlich und funktional unterscheidbaren Schularten in Deutschland wesentlich geprägt haben. Unterschieden werden zwei zentrale Entwicklungstendenzen: Zum einen die Entstehung des dreigliedrigen Bildungssystems bis zur Weimarer Republik (Abschnitt 2.1.1), zum anderen die Maßnahmen und Diskussionen um dessen Abschaffung bzw. Entschärfung im weiteren historischen Verlauf (Abschnitt 2.1.2).

2.1.1 Das dreigliedrige Bildungssystem entwickelt sich

Freiherr von Rochow forderte bereits 1779 in Preußen die verbindliche Einführung der Schulpflicht und gründete selbst Landschulen für Kinder aus Bauernfamilien. Rochow folgte damit dem Gedanken der Aufklärung, welche die „konkrete Beförderung des Bewusstseinswandels durch Bildung und Kultur“ (Meyer 2010, S. 14) in den Vordergrund stellte. Durch Bildung wollte man überholte Ideologien überwinden, den Geist befreien und Fortschritt fördern (Matthiessen 2000). Die Aufklärung kann damit als Gegenbewegung zum Absolutismus gesehen werden. Denn im Absolutismus sicherten Klerus und Adel durch Dogmatismus und die Berufung auf Gottes Gnaden ihre Vorherrschaft. Mit der Aufklärung setzte das Bürgertum gegen die Ständegesellschaft und die Willkürherrschaft des Klerus eine geistige Gegenbewegung. Den Geist befreien „aus seiner selbstverschuldeten Unmündigkeit“ (Kant 1784, S. 481) heißt, sich von Aberglaube und Dogmatismus zu lösen und stattdessen die Gegebenheiten oder auch naturwissenschaftlichen Phänomene rational zu durchdenken bzw. zu begründen. Bildung wird hier als Mittel der Emanzipation gesehen, um mit Wissen und Reflexionsvermögen die indoktrinierten und religiös begründeten Denkweisen zu überwinden. Die Aufklärung war überwiegend eine Bewegung des aufstrebenden Bürgertums, das am bisherigen politischen Gestaltungsprozess aufgrund der Ständeordnung nicht teilhaben durfte.

Ein gesellschaftlicher Wandel sollte jedoch durch die Bildung aller befördert werden.

Der Umsetzung der Schulpflicht und Bildung aller Kinder standen allerdings massive politische und soziale Barrieren im Wege. Das letzte Drittel des 18. Jahrhunderts war von Missernten und Massenarmut gekennzeichnet: „Im Hungerjahr 1770/1771 stieg der Brotpreis auf 12 Groschen, während der Lohn unverändert blieb. Der Lohn des Bergmanns reichte also nur für 2 Brote hin“ (Abel 1981, S. 49). In der Folge stieg die Sterblichkeitsrate drastisch an und ganze Bevölkerungsschichten verelendeten. Weitere finanzielle Belastungen durch den Bau von Schulen wären unter diesen Umständen für einen Großteil der Bevölkerung und für den Staat nicht tragbar gewesen. Zudem wurden die Kinder als Arbeitskräfte gebraucht (Diederich und Tenorth 1997; Fahrmeir 2017). Unter diesen Umständen konnte Rochow die Unterrichts- bzw. Schulpflicht nicht flächendeckend durchsetzen. Trotzdem hat er mit seinem Engagement und seinen modellhaften Landschulen die Entstehung eines institutionalisierten Bildungssystems in Deutschland maßgeblich vorangetrieben und Bildungsmöglichkeiten auch für Kinder jenseits des Adels geschaffen.

Etwa 30 Jahre nach Rochows Wirken war Wilhelm von Humboldt ab 1809 als Leiter der Sektion des Kultus und der Bildung damit beauftragt, das Konzept für ein liberales Bildungssystem zu entwerfen (Diederich und Tenorth 1997). „Liberal“ meint hier die Freiheit des Individuums, vornehmlich gegenüber staatlicher Regierungsgewalt, und richtet sich gegen Willkür, Machtmissbrauch und Staatsgläubigkeit (Giesinger 2016). Humboldt vertrat ein neuhumanistisches Bildungsideal. Im Gegensatz zu den Ideen der Aufklärung, die Wissen für wissenschaftlichen und gesellschaftlichen Fortschritt betont, soll Bildung zweckfrei sein. Die Individualität müsse ohne Rücksicht auf gesellschaftliche und aktuelle Bedürfnisse in der Schulerziehung ausgebildet werden. Die alten Sprachen gelten hierfür als zentrales Werkzeug (Konrad 2010). Der Entwurf seines Bildungssystems sah horizontale Strukturen vor, d.h., alle Kinder unabhängig des Standes sollten verschiedene Bildungsabschnitte durchlaufen können: Die Grundbildung, die Gymnasialbildung und das universitäre Studium. Je nach Interesse und Lebenslage sollten die Kinder die Schule, bspw. nach Abschluss der Grundbildung oder der Gymnasialbildung, verlassen (Diederich und Tenorth 1997).

Humboldts Vorstellungen von Schule und Lernen fielen in eine Zeit des Aufschwungs. Preußen war aus dem Heiligen Römischen Reich Deutscher Nation als starker Staat erwachsen. Die Gewerbefreiheit wurde eingeführt, die Bauern aus den Leibverhältnissen befreit, Menschenrechte eingeführt und eine parlamentarische Volksvertretung etabliert (Matthiessen 2000). Damit wurden die Unterschiede in Rechten und Privilegien zwischen den Ständen formal minimiert, was als Revolution von oben gesehen werden kann, welche ähnliche Entwicklungen wie in der Französischen Revolution vorbeugen sollte. Die Idee Humboldts eines horizontalen Bildungssystems wurde jedoch nicht realisiert. Stattdessen wurde die Differenz zwischen dem höheren und niederen Bildungswesen verstärkt. Dies kann als Reaktion der herrschenden Schicht auf den drohenden Machtverlust gesehen werden (Giesinger 2016; Konrad 2010).

Das meritokratische Berechtigungswesen verfestigte sich daher nur im höheren Bildungswesen: Der Zugang zu den Universitäten war nur noch mit Abitur möglich, wobei die Bildungsinhalte zunehmend extern definiert wurden (Drewek und Tenorth 2001). Diese Reformen stärkten die Stellung der berechtigenden höheren Schulen (Nath 2001) und somit auch die Stellung derer, die diese Schulen besuchten. Die inhaltliche Ausgestaltung wurde maßgeblich von Humboldts neuhumanistischem Bildungsideal geprägt, indem die Schwerpunkte in den alten Sprachen und der Philosophie lagen (Diederich und Tenorth 1997). Ziel war eine wissenschaftliche und klassische Bildung. Die höheren Schulen dienten damit der Herausbildung einer bildungsbürgerlichen Elite. Volksschulen hingegen erfüllten eine grundlegende Qualifikations- und Erziehungsfunktion, abzielend auf die „politische Integration und Disziplinierung der breiten Bevölkerung im Kontext der aufkommenden kapitalistischen Klassengesellschaft“ (Tenorth und Tippelt 2007, S. 644). Im Sinne des Berechtigungswesens hatten die Schulen des niederen Bildungswesens keine Selektionsfunktion (Nath 2001). Inhaltlich lag der Schwerpunkt in der Vermittlung grundlegender Fähigkeiten im Lesen und Rechnen sowie im Religionsunterricht (van Ackeren und Klemm 2011).

Die Struktur des Bildungssystems war somit zunächst überwiegend zweigliedrig (dualistisch) geprägt. Die Schularten (höheres und niederes Schulwesen) unterschieden sich dabei maßgeblich in den Berechtigungen bezogen auf die weitere Bildung bzw. den Beruf (Tenorth und Tippelt 2007). Durch die historischen Entwicklungen sowie die unterschiedlichen Funktionen und Ziele von höherem und niederem Schulwesen fanden sich weitere strukturelle Unter-

schiede, besonders in der Finanzierung und Ausstattung der Schulen. Höhere Schulen wurden durch Schulgelder, Kirchen und den Staat bezuschusst. Schulen des niederen Schulwesens waren durch die Gemeinden finanziert. Diese Zuwendungen waren deutlich geringer als an Schulen der höheren Bildung. Auch das Lehrpersonal unterschied sich zwischen den beiden Schularten. Die Lehrenden an höheren Schulen waren studierte Fachlehrer im höheren Beamtendienst. An niederen Schulen arbeiteten die Lehrenden hingegen als Klassenlehrer (d.h. fächerübergreifend). Sie verfügten als mittlere Beamte über eine geringere Eigenbildung (Diederich und Tenorth 1997).

Das mittlere Schulwesen und somit das dreigliedrige Bildungssystem etablierte sich erst im weiteren gesellschaftlichen Wandel. In den Jahren um 1850 bis 1880 stieg die Wirtschaft deutlich an. Umstrukturierungsprozesse in Landwirtschaft, Handel und Gewerbe kündigen die aufkommende Industrialisierung an. Das Deutsche Kaiserreich hatte sich von der Agrar- zur Industrienation entwickelt. Die Menschen zogen vom Land in die Städte, um in den neu aufkommenden Berufen zu arbeiten. Im Zuge dessen bildete sich eine breite Mittelschicht, die überwiegend in Industrie, Handel und Gewerbe tätig war (Matthiessen 2000). Bildung wurde mit wachsenden Fortschritten in Medizin und Technik inzwischen weniger im Sinne der Individualität als vielmehr funktional gesehen (Drewiek und Tenorth 2001). Somit wurde mit den wachsenden Anforderungen in den Berufen auch eine Anpassung des Bildungssystems als notwendig empfunden (Drewiek 2015). Zudem forderte die gewachsene Mittelschicht eine höhere Bildungsbeteiligung ein (Völk 2014), was sich auch in steigenden Abiturientenzahlen zeigte (Drewiek und Tenorth 2001). Als Reaktion darauf etablierte sich das mittlere Schulwesen, welches inhaltlich stärker auf technisch naturwissenschaftliche Fächer fokussierte, sogenannte Realien (Diederich und Tenorth 1997). Hier finden sich auch Schulen, welche ehemals das höhere Schulwesen bedienten, aber die Auflagen für die Vergabe des Abiturs nicht erfüllten. Vorrangiges Ziel der mittleren Schulen war die Vorbereitung auf bestimmte berufliche Funktionen, ohne jedoch den Charakter als allgemeinbildende Schule einzubüßen (Wollenweber 1997). Zwar passten sich auch die Gymnasien inhaltlich an die geänderten Anforderungen an, indem auch hier Schulen mit einem stärkeren Fokus auf naturwissenschaftliche und technische Fächer als Schulen mit Abiturrecht in die höhere Bildung aufgenommen wurden (Tenorth und Tippelt 2007). Es blieb jedoch bei der strikten Trennung, dass höhere Schulen mit einem Schwerpunkt im Reflexionswissen auf die universitäre Bildung und spätere Führungsaufgaben vorbereiteten. Mittlere Schulen

bereiteten auf die Arbeit in komplexeren Berufen in Industrie und Handel vor. Das niedere Schulwesen vermittelte eine Grundbildung mit einem Schwerpunkt auf Rechnen, Schreiben, Religion und Heimatkunde. Diese Unterschiede spiegeln sich weiterhin im Fächerkanon und den Stundentafeln der Schularten wider. Eine Gegenüberstellung der Unterrichtsinhalte findet sich bspw. bei Dierich und Tenorth (1997).

Weiter verfestigt wurde das entstandene dreigliedrige Schulsystem in Zuge der staatlichen Neuorganisierung nach dem Ersten Weltkrieg. Während des Wiederaufbaus und der Ausgestaltung des Staates zur demokratischeren Weimarer Republik wurden gesellschaftliche Reformen umgesetzt, welche stärkere Partizipation und Gleichberechtigung verankerten (Fahrmeir 2017). Auch das Bildungssystem wurde im Zuge dessen hinsichtlich der Dreigliedrigkeit diskutiert. Befürworter einer Einheitsschule konnten sich in den Reformdebatten der Weimarer Republik jedoch nicht durchsetzen. Als eine Art Kompromiss wurde die gemeinsame vierjährige Grundschule eingeführt. Daran schließen sich die drei Schularten Gymnasium, Realschule und Hauptschule³ an (Tenorth und Tippelt 2007). Die inhaltliche und funktionale Abgrenzung der Schularten blieb bestehen bzw. wurde sogar verstärkt, indem die Lehrpläne für Schulen des mittleren Bildungswesens stärker auf die beruflichen Laufbahnen ihrer Lernenden zugeschnitten wurden. Fokussierungen gab es auf die Bereiche Handel und Verkehr, Gewerbe, Hauswirtschaft sowie sozial-pflegerische Berufe (Wollenweber 1997). In dieser Form ist das dreigliedrige Bildungssystem heute noch bekannt, oft bezeichnet als „traditionelles Bildungssystem“ in Abgrenzung zum Gesamtschulsystem.

3 Im Folgenden werden die Schularten des traditionellen, dreigliedrigen Bildungssystems benannt als Gymnasium, Realschule und Hauptschule. Die Bezeichnung Gymnasium für Schulen, welche sich aus dem höheren Bildungswesen entwickelt haben und das Abitur vergeben, ist in der Weimarer Republik flächendeckend etabliert und bis heute gültig. Realschulen werden zur Zeit der Weimarer Republik teilweise noch als Mittelschulen bezeichnet. Die Schulen des niederen Schulwesens werden in der Weimarer Republik noch als Volksoberschulen betitelt und ab den 50er Jahren überwiegend als Hauptschulen bezeichnet. Eine Auflistung der Schularten und Bezeichnungen in den einzelnen Bundesländern findet sich im Anhang 12.1 (Stand 2018).

2.1.2 Das dreigliedrige Bildungssystem im Wandel

Im historischen Verlauf von der Umsetzung der Unterrichtspflicht bis zum verstetigten dreigliedrigen System in der Weimarer Republik hat sich das Bildungssystem schrittweise der breiten Bevölkerung geöffnet. Parallel wurden die Bildungsziele der Schularten konkretisiert und differenziert. Hervorzuheben ist, dass mit steigender Bildungsbeteiligung und inhaltlichen Reformbestrebungen tendenziell neue Schularten gegründet wurden, statt die Zielgruppe der bestehenden Schulart(en) zu erweitern und die Inhalte zu flexibilisieren. Ähnliche Tendenzen zeigen sich auch im weiteren historischen Verlauf des Bildungssystems in Deutschland: Kritik und Reformbestrebungen bezüglich des dreigliedrigen Bildungssystems werden überwiegend mit der Begründung paralleler Schularten – integrierte Gesamtschulen und Schulen mit mehreren Bildungsgängen – abgefangen. Diese Entwicklungen des dreigliedrigen Systems hin zum vier- bzw. fünfgliedrigen werden im Folgenden näher ausgeführt.

Nach dem zweiten Weltkrieg war die Struktur des Bildungssystems in Deutschland infolge der politischen Konsequenzen einem Wandel und der Diskussionen unterworfen. Die Entwicklungen vollzogen sich in Ost- und Westdeutschland jedoch zunächst gegenläufig. In der sozialistisch geprägten Deutschen Demokratischen Republik (DDR) wurde entsprechend dem Gleichheitsgedanken ein Einheitsschulsystem aufgebaut. Alle Kinder wurden von der ersten bis zur zehnten Klasse gemeinsam unterrichtet (Baumert et al. 2008; Tillmann et al. 2008). Hier reduziert sich das ehemals dreigliedrige Schulsystem zunächst auf eine Schulart. Die meisten Lernenden mündeten im Anschluss in die Berufsbildung ein, während ausgewählten Lernenden der Zugang zu Abitur und Hochschule ermöglicht wurde (van Ackeren und Klemm 2011). Dieser war jedoch nicht rein über akademische Leistungen geregelt, sondern auch durch politisches Wohlverhalten. Das Einheitssystem wurde in der Bundesrepublik Deutschland (BRD) deshalb als rückständig verurteilt und hatte auf Dauer keinen Bestand (Baumert et al. 2008).

In der BRD wurde die traditionelle dreigliedrige Struktur zunächst beibehalten, trotz Plädoyer der Besatzungsmächte für ein demokratisches, inklusives Bildungssystem (Tillmann et al. 2008). Gerechtfertigt wurde die Gliederung des Bildungssystems durch eine nativistische Begabungstypologie und eine dieser Typologie entsprechende Klassifikation von Berufsgruppen. „Die Hauptschule galt als Schule der ‚praktisch Begabten unteren Volksschichten‘, für die ein

ganzheitlicher, anschaulicher und erlebnisnaher Unterricht, der das Heimatprinzip betonte, angemessen erschien. Im Gymnasium dagegen sollten die ‚abstrakt Begabten‘ auf ein Hochschulstudium und gesellschaftliche Führungspositionen vorbereitet werden.“ (Baumert et al. 2008). Die Realschulen waren hingegen an der Schnittstelle als Schulen für praktisch-theoretisch Begabte angedacht mit einer deutlichen Zuordnung zu Berufsfunktionen (Wollenweber 1997). „Sie [die Realschule] bereitet ihre Schüler auf Aufgaben des praktischen Lebens mit erhöhter fachlicher, wirtschaftlicher und sozialer Verantwortung vor und vermittelt die dafür notwendige allgemeine Bildung.“ (KMK 2014, S. 150)⁴

Analysen des Bildungssystems der BRD offenbarten in den 1960er Jahren jedoch große soziale, regionale und geschlechtsspezifische Ungleichheiten in der Bildungsbeteiligung, veranschaulicht im Bild der „katholischen Arbeiter-tochter vom Lande“ (Peisert 1967). Diese Zuspitzung hat dazu geführt, dass die begabungstypologische Verteilung auf die Schularten zunehmend in Frage gestellt wurde. Das Bildungssystem durchlief daraufhin einen Modernisierungsprozess, welcher sich Ende der 1960er Jahre in einem ideologischen Streit der Bundesländer bzw. der Parteien über die Strukturfragen des Bildungssystems verding und bis weit in die 1980er Jahre andauerte (vgl. Tillmann 1987, 1996). Politisch standen sich zwei Lager gegenüber. Starke Reformkräfte (SPD, Gewerkschaften, Teile der FDP) plädierten für einen Wandel des gegliederten Bildungssystems hin zu integrierten Gesamtschulen. Mit integrierten Schulen möchte man soziale Integration erreichen und die Chancengleichheit erhöhen. Zudem könne Heterogenität als Lerngelegenheit genutzt und individuelle Profile gestärkt werden. Auch eine geringere Weichenstellung und demografische Veränderungen wurden seinerzeit bereits diskutiert. Eher konservative Kräfte (CDU, Unternehmerverbände, katholische Kirche) wollten das gegliederte Schulsystem beibehalten und intern modernisieren. Hier wird die Position vertreten, dass das dreigliedrige Schulsystem optimale Förderung nach Interessen und Fähigkeiten biete. Zudem könne an Gymnasien die Bildungselite optimal gefördert werden (Tillmann et al. 2008).

Ende der 1960er Jahre einigten sich die Länder auf einen Gesamtschulversuch. Fend und Kollegen führten bspw. in SPD-regierten Bundesländern eine großangelegte – und im Kontext der Gesamtschuldiskussion meist zitierte – Begleit-

4 Die Formulierung geht auf einen Beschluss der KMK von 1953 zurück, siehe hierzu Maskus (1966, S. 141).

studie durch (Fend 1982). Fend und Kollegen folgerten, dass „bei gleichem Begabungsniveau [...] die Chancen, auf höherem, dem Gymnasialniveau äquivalenten Kursniveau zu sein, von der sozialen Herkunft unabhängig [waren]“ (Fend 2009, S. 40). Die positiven Ergebnisse wurden jedoch durch viele Probleme überschattet: „Die verbesserte Chancengerechtigkeit war in vielen Gesamtschulen [...] von einer Leistungsminderung begleitet. Auch die erzieherischen Ergebnisse konnten nicht durchwegs überzeugen“ (Fend 2009, S. 40). Zudem unterschieden sich die Modellschulen und die Auslegung der Ergebnisse stark zwischen den Bundesländern bzw. zwischen den regierenden Parteien. Auch in der breiten Bevölkerung herrschte keine Einigkeit über die Einführung von Gesamtschulen. In der Konsequenz wurden in einigen Bundesländern Gesamtschulen zusätzlich zum traditionellen Bildungssystem etabliert (bspw. Hessen, Niedersachsen, Hamburg, Nordrhein-Westfalen). Andere Bundesländer kehrten zum traditionellen dreigliedrigen Bildungssystem zurück (bspw. Bayern, Baden-Württemberg, Rheinland-Pfalz). Das ursprüngliche Ziel – das gegliederte Schulsystem durch die Gesamtschulen zu ersetzen – konnte jedoch in keinem Bundesland umgesetzt werden (Baumert et al. 2008). Stattdessen hat sich das Bildungsangebot in der Diskussion um Bildungsbeteiligung ein weiteres Mal erweitert. Die traditionellen Schularten – Hauptschule, Realschule, Gymnasium – blieben bestehen.

Im weiteren Verlauf der Diskussion um das Bildungssystem rückten Fragen zur Gliederung des Bildungssystems in den Hintergrund. Stattdessen lag der Fokus bei der bestmöglichen Qualifikation der Lernenden für die gestiegenen beruflichen Ansprüche sowie bei der Durchlässigkeit zwischen den Schularten (Dühlmeier 2009). Auch im Zuge der Wiedervereinigung (1989/90) hat das traditionelle Schulsystem Bestand. Im Aushandlungsprozess um ein deutschlandweites einheitliches Schulsystem wurde die dominante Stellung der BRD sowie die politische Abgrenzung zum Kommunismus deutlich (Fahrmeir 2017). Die neugegründeten Bundesländer der ehemaligen DDR sollten ungeachtet der Diskussion um die Gesamtschulen in den 1960er und 1970er Jahren das Einheitsschulsystem abschaffen und das traditionelle dreigliedrige Schulsystem übernehmen (van Ackeren und Klemm 2011). Den Vorgaben standen jedoch ideologische sowie demografische Barrieren gegenüber, bspw. war es durch geringe und sinkende Bevölkerungszahlen in weiten Teilen der neuen Bundesländer schwer möglich, drei Schularten anzubieten. Umgesetzt wurde in den neuen Bundesländern letztlich überwiegend ein zweigliedriges Schulsystem. Dem Gymnasium wurden kombinierte Haupt- und Realschulen gegenüberge-

stellt (Baumert et al. 2008). Diese Schulen sind in die Kategorie der Schulen mit mehreren Bildungsgängen einzuordnen, allerdings mit der Besonderheit nicht alle Bildungsgänge anzubieten wie die Gesamtschulen der alten Bundesländer, sondern zwei. Zusammenfassend verdeutlichen auch die Entwicklungen im Zuge der Wiedervereinigung die Stabilität des dreigliedrigen Systems (in den alten Bundesländern), die exklusive Stellung des Gymnasiums (in den alten und neuen Bundesländern) sowie die Tendenz, die Bildungslandschaft um neue Schularten zu erweitern.

2.2 Das Bildungssystem der Sekundarstufe I heute

Das Bildungssystem in Deutschland befindet sich weiter im Wandel, die Strukturdebatte ist jedoch auf nationaler Ebene in den Hintergrund geraten. Die international vergleichenden Schulleistungsstudien seit dem Jahr 2000 haben auf politischer Ebene eine intensive Debatte um die Qualität und die Fortentwicklung des Schulsystems ausgelöst und auch die Nutzer des Bildungssystems weiter für bildungsstrategische Überlegungen sensibilisiert (Hurrelmann 2013; Liegmann 2016). Strukturfragen bezogen auf die Gliederung des Bildungssystems werden auf nationaler Ebene vermieden und somit auch die Differenzen zwischen den Schularten in ihren Zielen und Inhalten wenig reflektiert (Tillmann et al. 2008). Stärker im Vordergrund stehen die Handlungsfreiheit der Einzelschule sowie die Qualitätssteuerung, unter anderem über Bildungsstandards. In der Konsequenz zeigt sich überwiegend eine stillschweigende Änderung der Dreigliedrigkeit durch verändertes Schulwahlverhalten und durch den demographischen Wandel. Der Rückgang der Schülerzahlen erfordert besonders in Kleinstädten und im ländlichen Raum, dass Schulen bzw. Schularten zusammengelegt werden (Zymek 2013). Nur innerhalb weniger Bundesländer – vorwiegend in den nördlich gelegenen Bundesländern – wurde die Strukturfrage angegangen und Pilotprojekte angestoßen bzw. das Bildungssystem umstrukturiert. In diesen Bundesländern zeigt sich ein Trend zur Zweigliedrigkeit (Autorengruppe Bildungsberichterstattung 2018).

Welche Konsequenzen die politischen Linien sowie die gesellschaftliche Nachfrage auf die Struktur des Bildungssystems weiterhin haben werden, bleibt eine spannende Frage und wird wohl auch von langfristigen Erfahrungen und Konsequenzen der strukturellen Änderungen in den nördlich gelegenen Bundesländern beeinflusst sein. Aktuell kann aus nationaler Perspektive noch ein Bil-

dungssystem mit fünf übergreifenden Schularten mit unterschiedlichen Inhalten bzw. Leitbildern identifiziert werden, welche im Anschluss an die Grundschule besucht werden. Aus dem traditionellen dreigliedrigen System sind bis heute *Hauptschulen (HS)*, *Realschulen (RS)* und *Gymnasien (GY)* weiterhin Bestandteile des deutschen Bildungssystems. Die Abgrenzung dieser drei Schularten in ihren Inhalte und Ziele wurde seit der Weimarer Republik nur wenig überarbeitet. Wie gezeigt wurde, haben sich die weiteren Debatten über die Strukturfragen in den 1970er Jahren und im Zuge der Wiedervereinigung in der Gründung neuer, bildungsgangübergreifender Schularten entladen, welche somit ideologisch andere Schwerpunkte verfolgen. In den *integrierten Gesamtschulen (IGS)* werden die Lernenden gemeinsam unterrichtet. Der Unterricht in den Hauptfächern findet jedoch in Kursen mit (meist drei) unterschiedlichen Niveaustufen in Anlehnung an die Bildungsgänge statt. Alle übrigen Schulen mit zwei oder drei Bildungsgängen werden in dieser Arbeit unter die Kategorie der *Schulen mit mehreren Bildungsgängen (MBG)* gefasst. Die Definition der bildungsgangübergreifenden Schulen erfolgt ausführlicher in Abschnitt 2.2.4.

Die KMK kontrastiert in der „Vereinbarung über die Schularten und Bildungsgänge im Sekundarbereich I“ (KMK 2006) die Schularten des dreigliedrigen Bildungssystems entsprechend der angedachten Bildungsziele und Bildungswege (Abbildung 1). Hier finden sich deutliche Abstufungen in den Bildungsmöglichkeiten und Anschlussperspektiven. Bildungsgangübergreifende Schulen – hier differenziert nach Schularten mit mehreren Bildungsgängen und integrierten Gesamtschulen – definiert die KMK hingegen über die Organisationsform des leistungsdifferenzierten Unterrichts. Bildungsziele und Bildungswege werden für die Beschreibung dieser Schularten nicht berücksichtigt (KMK 2006).

		Bildungsziel	Bildungsweg
Traditionelles Bildungssystem	Hauptschule	„grundlegende allgemeine Bildung“	„vor allem in berufs-, aber auch in studienqualifizierenden Bildungsgängen“
	Realschule	„erweiterte allgemeine Bildung“	„in berufs- und studienqualifizierenden Bildungsgängen“
	Gymnasium	„vertiefte allgemeine Bildung“	„an einer Hochschule, aber auch in berufsqualifizierenden Bildungsgängen“
Bildungsgang-übergreifende Schulen		Organisationsform	
	IGS	„der Unterricht [wird] entweder in abschlussbezogenen Klassen oder – in einem Teil der Fächer – leistungsdifferenziert auf mindestens zwei lehrplanbezogen definierten Anspruchsebenen in Kursen erteilt.“	
	MBG		

Abbildung 1. Gegenüberstellung der Schulartbeschreibungen.

In Anlehnung an KMK (2006, S. 7).

Ziel der folgenden Abschnitte ist es, die Unterschiede in den Schularten, welche sich aus den historischen Entwicklungsprozessen ergeben haben, detaillierter herauszuarbeiten. Bildungsziele und -inhalte werden ebenso berücksichtigt wie die Zugangsvoraussetzungen, die Schülerzusammensetzung und Anschlussperspektiven.

2.2.1 Hauptschule

Die Hauptschule (in Bayern inzwischen Mittelschule genannt) beerbt aus der historischen Entwicklung heraus Schulen des niederen Schulwesens und der späteren Volksoberschulen (Leschinsky 2008a). War dies einst die Schulart mit dem größten Anteil an Lernenden (1970 besuchten noch 52,7 % der Lernenden eine Hauptschule; Ipfling 2009), ist die Hauptschule inzwischen als eigenständige Schulart auf dem Rückzug (Diederich und Tenorth 1997). Viele Jahre musste sie als Regelschule in den meisten Bundesländern obligatorisch angeboten werden. Schulpflichtige Kinder, die keine andere Vollzeitschule oder entsprechende staatlich anerkannte Ersatzschule besuchen, sind zum Besuch der Hauptschule (oder dem Hauptschulzweig) verpflichtet. Für die Aufnahme in die Hauptschule besteht daher keine leistungsbezogene Eingangsvoraussetzung (Ipfling 2009). Lediglich eine Ausgrenzung in das Förderschulwesen aufgrund von Förderbedarfen, welche nicht an allgemeinen Schulen bedient werden können, ist möglich (Bundschuh und Baier 2009; KMK 1994, 2016b). Im allgemeinen mehrgliedrigen Bildungssystem wird die Hauptschule deshalb als die Schulart mit dem geringsten Anspruch angesehen (Leschinsky 2008a). Auch

eine praktische Eingangsprüfung oder Auslese wird nicht getroffen, obwohl die Hauptschule oft als die „Schule für die praktisch Begabten“ bezeichnet wurde (Ipfling 2009). Eltern, deren Kinder die Hauptschule besuchen, haben durchaus positive Erwartungen an diese Schulart. Sie sehen die Hauptschule oft nicht als Bildungsziel, sondern als Bildungsweg. Von der Schule und den Lehrkräften wird erwartet, dass diese besonders auf die Bedürfnisse und Lernvoraussetzungen der Kinder eingehen, so dass sich die Lernleistungen der Kinder stabilisieren und auch weiterführende Bildungsabschlüsse nicht ausgeschlossen sind (Rekus et al. 1998).

Die gegenwärtigen Hauptschulen erfüllen diese Erwartungen noch nicht. Das erklärte Bildungsziel dieser Schulart ist weniger die Vorbereitung auf höhere Bildungsabschlüsse als vielmehr die berufspraktische Vorbereitung in eingeschränkten Berufsfeldern. Die Allgemeinbildung ist auf grundlegendes Wissen beschränkt (Ipfling 2009). Diese Ziele wurden wesentlich im Rahmenplan zur Umgestaltung und Vereinheitlichung des allgemeinen öffentlichen Schulwesens vom Deutschen Ausschuss (1957) festgelegt: „[Die Hauptschule] als Schule, die den Zusammenhang von Allgemeinbildung und Berufsleben herstellt“ (Ipfling 2009, S. 239). Inhaltlich hat die Hauptschule daher einen starken Fokus auf praxisorientierte Unterrichtsprojekte. Diese sollen in Verbindung mit Praktika langfristige Perspektiven für die Lernenden durch Kontakte und Praxiserfahrungen anbahnen (Leschinsky 2008a). Kooperationen sind vorrangig mit Betrieben der „niedrigen Arbeitswelt“ geschlossen, bspw. im Handwerk oder in haushaltsnahen Bereichen (Ipfling 2009). Ergänzt wird die praxisnahe Ausbildung durch Unterrichtsfächer wie Arbeitslehre und informationstechnische Wissensvermittlung. Zur allgemeinen Grundbildung gehört neben klassischen Fächern wie Deutsch und Mathematik inzwischen auch der fremdsprachliche Unterricht in Englisch.

Besucht werden Hauptschulen überwiegend von Lernenden aus bildungsfernen Familien und von Lernenden mit Migrationshintergrund. Der Anteil an Jugendlichen aus Arbeiterfamilien beträgt ca. 60 % (Nold 2010); der Anteil von Lernenden mit Migrationshintergrund beläuft sich auf ca. 51 % (Statistisches Bundesamt 2017a). Die Zusammensetzung der Schülerschaft ist entgegen der angedachten leistungsbezogenen Selektion des Deutschen Bildungssystems nicht rein auf die Leistung zurückzuführen. So finden Baumert et al. (2006b) deutliche Überschneidungen in den Leistungsbereichen der Lernenden zwischen den verschiedenen Schularten. Vielmehr wirken herkunftsbedingte Selektionspro-

zesse während des Übergangs von der Grundschule in die Sekundarstufe (Maaz et al. 2010).

Die Hauptschule endet mit der neunten bzw. zehnten Klasse (KMK 2017a) und entlässt ihre Lernenden in das Berufsbildungssystem, entweder direkt in die praktische Berufsausbildung oder in berufliche Qualifizierungsmaßnahmen im Übergangssystem. Unter den Sammelbegriff „Übergangssystem“ fallen Bildungsangebote, welche verpflichtend sind für Abgänger des allgemeinen Bildungssystems, die ihre Pflichtschulzeit noch nicht beendet haben und nicht in die berufliche Ausbildung einmünden (bspw. Berufsgrundbildungsjahr oder das Berufsvorbereitungsjahr). Zudem gibt es Anschlussmöglichkeiten zum Erwerb zusätzlicher Bildungsqualifikationen bspw. in Form eines zusätzlichen Schuljahres zum Erwerb des Mittleren Abschlusses (Leschinsky 2008a). Trotz des berufspraktischen Schwerpunktes der Hauptschulbildung ist der Anteil von Lernenden mit Hauptschulabschluss in der Dualen Bildung oder des Schulberufssystems vergleichsweise gering. Im Jahr 2014 haben 26,5 % der Jugendlichen in der Dualen Bildung und 17,4 % der Jugendlichen in beruflichen Vollzeitschulen einen Hauptschulabschluss, mit sinkender Tendenz. Realschüler sind hingegen mit 48,3 % bzw. 59,7 % in der Berufsbildung vertreten. In den Ausbildungsbereichen Hauswirtschaft (51,7 %) Handwerk (46,2 %) und Landwirtschaft (32,9 %) sind Jugendliche mit Hauptschulabschluss am häufigsten vertreten. Im öffentlichen Dienst finden nur 3,7 % der Jugendlichen mit Hauptschulabschluss einen Ausbildungsplatz (Bundesinstitut für Berufsbildung 2016). Auch eine Analyse des Deutschen Gewerkschaftsbundes macht deutlich, dass Jugendlichen mit Hauptschulabschluss der Ausbildungsmarkt nur begrenzt zugänglich ist. So waren im Juli 2016 lediglich 39,05 % der angebotenen Ausbildungsplätze für Jugendliche mit Hauptschulabschluss ausgeschrieben (Anbuhl 2016). Den größeren Anteil stellen Jugendliche aus Hauptschulen im Übergangssystem. Zudem steigt der Anteil an Jugendlichen, welche einen höheren schulischen Bildungsabschluss anstreben (Autorengruppe Bildungsberichterstattung 2016).

2.2.2 Realschule

Die Realschule beerbt die Schulen der mittleren Schulbildung. Soziale und wirtschaftliche Interessen haben seit Begründung der preußischen Mittelschulen zur verstärkten Nachfrage dieser Schulart bis ins Jahr 2000 beigetragen und

die inhaltliche Ausgestaltung geprägt (Leschinsky 2008b). Ähnlich der Hauptschule befindet sich die Realschule als eigenständige Schulart auf dem Rückzug (Völk 2014). Heute haben Gymnasien und kombinierte Schularten den größten Zulauf. Trotzdem besuchen im Schuljahr 2014/2015 noch fast 20 % der Lernenden in Deutschland eine Realschule (Autorengruppe Bildungsberichterstattung 2016). Die Realschule kann direkt im Anschluss an die Grundschule besucht werden. Die Lernenden müssen hierfür einen bestimmten Notendurchschnitt in der vierten Klasse vorweisen bzw. eine entsprechende Empfehlung aus der Grundschule erhalten oder den Probeunterricht bzw. die Orientierungsstufe erfolgreich absolvieren.⁵ Weiterhin ist mit steigender Durchlässigkeit zwischen den Schularten und Bildungsgängen ein Wechsel aus einer anderen Schulart möglich. Abwärts gerichtete Durchlässigkeit stellt keine Mindestanforderungen an die schulischen Leistungen der Lernenden. Aufwärtsgerichtete Durchlässigkeit ist hingegen mit einem bestimmten Notendurchschnitt und/oder erfolgreichem Probeunterricht verbunden.

Bildungstheoretische und politische Leitbilder haben die inhaltliche Ausgestaltung der Realschulen weniger geprägt. Vielmehr geht das Bildungsprogramm auf die Nachfrage in der breiten Bevölkerung und Anforderungen der Wirtschaft zurück (Leschinsky 2008b). Ziel ist daher, Allgemeinbildung mit einem Bezug zum praktischen Leben zu vermitteln. Der Unterricht der Realschule soll problemorientiert erfolgen, anhand konkreter Gegenstandsbereiche. Die theoretischen Grundlagen dienen der Durchdringung von Sachverhalten, verfolgen aber keinen Eigenzweck, wie etwa im gymnasialen Unterricht (Fees 2009). Durch die Berufsbezogenheit im Bildungsverständnis hebt sich die Realschule didaktisch gegenüber dem Gymnasium ab. Im Vergleich zur Hauptschule wird die Allgemeinbildung durch ein umfassenderes fremdsprachliches Angebot sowie ein breiteres mathematisch-naturwissenschaftliches Lehrangebot erweitert (Leschinsky 2008b). Als Besonderheit der Realschule ist die innere Differenzierung nach Interessen zu sehen. Die Lernenden an Realschulen können hierfür (je nach Schule) aus drei bis vier Wahlpflichtfächergruppen einen Schwerpunkt auswählen: hauswirtschaftlich-sozial, wirtschaftlich, sprachlich oder mathematisch-technisch-naturwissenschaftlich (Völk 2014). Die Schwerpunkte sind unabhängig von den Vorleistungen der Lernenden wählbar und

5 Eine Zusammenstellung der Übergangsregeln der einzelnen Länder fasst der Bildungsbericht (2016) in Tabelle D2-5web zusammen.

bereiten auf die vielfältigen Anschlussmöglichkeiten in der beruflichen sowie schulischen Bildung vor (Wollenweber 1997).

Innerhalb des traditionellen Bildungssystems waren und sind Realschulen vergleichsweise offen gegenüber allen sozialen Gruppen der Bevölkerung (Wollenweber 1997). Im Jahr 2012 wechselten 30 % der Lernenden mit niedrigem sozioökonomischem Status von der Grundschule in eine Realschule. Bei den Lernenden mit mittlerem und hohem sozioökonomischem Status lag der Anteil bei 32 % bzw. 17 %.⁶ Es zeigt sich jedoch über einen Zeitraum von 12 Jahren, dass Lernende mit mittlerem und hohem sozioökonomischen Status mit wachsendem Anteil auf Gymnasien wechseln (Autorengruppe Bildungsberichterstattung 2016). Hier spiegeln sich vermutlich die geänderten Bildungsaspirationen der Lernenden bzw. ihrer Eltern wider, welche sich vermehrt auf das Abitur bzw. die Hochschulbildung richten.

Die Realschule endet nach der 10. Klasse mit dem Mittleren Abschluss bzw. der Mittleren Reife (KMK 2017a) und öffnet ein breites Spektrum an Möglichkeiten für den weiteren Bildungs- und Berufsweg (Wollenweber 1997). Ein großer Teil der Lernenden mündet in die duale oder vollschulzeitliche Berufsbildung, vorwiegend in anspruchsvollen Berufen im öffentlichen Dienst oder in Industrie und Handel (Bundesinstitut für Berufsbildung 2016). Zudem berechtigt der Mittlere Abschluss zum Besuch weiterführender schulischer Einrichtungen in der Sekundarstufe II, welche mit wachsendem Interesse nachgefragt werden. Je nach Notendurchschnitt und Region können die Lernenden auf ein (Wirtschafts-/berufliches) Gymnasium, eine Gesamtschule mit gymnasialer Oberstufe, eine Berufsoberschule oder eine Fachhochschule wechseln (KMK 2017a). Zudem stehen Weiterqualifizierungsmaßnahmen im Übergangssystem, wie das Berufsgrundschuljahr, offen.

6 Der Bildungsbericht bildet die Kategorien für den höheren, den mittleren und den niedrigen sozialen Status anhand des Index für den höchsten beruflichen Status der Eltern (HISEI). Gegenübergestellt werden die 25 % der Lernenden mit den höchsten Indexwerten (Hoch), diejenigen 50 % mit mittleren (Mittel) und jene 25 % mit den niedrigsten Indexwerten (Niedrig).

2.2.3 Gymnasium

Das Gymnasium ist die Schulart mit der längsten Tradition und höchsten Akzeptanz in der breiten Bevölkerung. Bis heute existiert diese Schulart in allen Bundesländern (KMK 2017a) und weist trotz verschiedener Reformbestrebungen hohe Konsistenz in Inhalten und Zielen auf (Haag und Hopperdietzel 2009). Das Gymnasium kann direkt im Anschluss an die Grundschule besucht werden (KMK 2017a). Die Lernenden müssen hierfür einen bestimmten Notendurchschnitt im letzten Grundschuljahr vorweisen bzw. eine entsprechende Empfehlung aus der Grundschule erhalten oder den Probeunterricht bzw. die Orientierungsstufe erfolgreich absolvieren.⁷ Weiterhin ist mit steigender Durchlässigkeit zwischen den Schularten und Bildungsgängen ein Wechsel aus einer anderen Schulart möglich (KMK 2017a). Trotz – vielleicht auch wegen – ihres Rufes als elitäre und exklusive Bildungsanstalt steigt der Anteil der Lernenden, welche nach der Grundschule auf ein Gymnasium wechseln, stetig an, von ehemals 3 % in den 1930er Jahren über 29 % im Jahr 2005 (Trautwein und Neumann 2008) hin zu 40 % im Schuljahr 2014/2015 (Statistisches Bundesamt 2016). Dadurch ist das Gymnasium vielerorts von der Schulart der Elite zur Schulart für die meisten geworden, weshalb das Gymnasium in der bildungspolitischen Diskussion oft als „neue Hauptschule“ tituliert wird (Carvalho 2017).

Gymnasien haben die doppelte Bestimmung, sowohl eine vertiefende Allgemeinbildung zu vermitteln als auch Studierfähigkeit anzubahnen. Realisiert wird dies über Pflicht- und Wahlfächer sowie Grund- und Leistungskurse (in der Oberstufe), deren Fächerkanon sich zu einem großen Teil aus universitären Wissenschaften ableitet (Kiper 2007b; Krimm 2009). Ziel ist es, die Lernenden zur Hochschulreife zu führen. Die Unterrichtsinhalte haben demzufolge einen viel stärkeren wissenschaftspropädeutischen Charakter als die anderer Schularten der Sekundarstufe. Zudem nimmt die kulturelle Bildung einen wesentlichen Teil ein. Neben den alten Sprachen finden sich – vor allem im Nachmittagsangebot – vielfältige Möglichkeiten zur künstlerischen und musischen Aktivität (Haag und Hopperdietzel 2009). Mit den steigenden Schülerzahlen aber auch Forderungen aus der Wirtschaft erhöht sich der Druck auf Gymnasien, nicht nur kulturell und wissenschaftspropädeutisch sondern auch berufsvorbereitend auszubilden (Kiper 2007a). Demgegenüber sind die Gymnasien jedoch ver-

7 Eine Zusammenstellung der Übergangsregeln der einzelnen Länder fasst der Bildungsbericht (2016) in Tabelle D2-5web zusammen.

gleichsweise resistent. Wie keine andere Schulart betont das Gymnasium den fachlichen Zugang zu den Gegenständen sowie eine starke Reflexionskultur. In den Hintergrund treten an dieser Schulart hingegen das Erfahrungslernen sowie erzieherische Ansprüche. Auch verschiedene Anstöße auf wissenschaftlicher, politischer und administrativer Ebene zu einer nachhaltigen Reform von Curriculum und Unterrichtsformen erreichen die unterrichtliche Gestaltung nur gering (Trautwein und Neumann 2008).

Mit dem Anstieg der Schulbesuchsquoten an Gymnasien vollzog sich auch eine Öffnung für bisher benachteiligte soziale Schichten. Damit ist das Gymnasium bezogen auf die soziale Herkunft inzwischen die heterogenste Schulart (Kiper 2007a). Trotzdem bleibt das Gymnasium sehr selektiv. So sind die relativen Unterschiede in der gymnasialen Beteiligung zwischen den sozialen Schichten weitgehend konstant geblieben (Trautwein und Neumann 2008). Im Jahr 2012 besuchen 69 % der 15-Jährigen mit hohem sozioökonomischem Status ein Gymnasium. Bei den 15-Jährigen mit mittlerem und niedrigem sozialem Status sind es nur 36 % bzw. 15 % (Autorengruppe Bildungsberichterstattung 2016). Insgesamt ist damit der sozioökonomische Status der Eltern höher als an anderen Schularten und der Anteil an Lernenden mit Migrationshintergrund relativ gering (Kiper 2007a). Zudem nimmt das Gymnasium nach wie vor eine Schülerschaft auf, die intellektuell deutlich positiv selektiert ist (Köller 2007; Prenzel et al. 2013). Die vergleichsweise hohe Abwärtsmobilität verstärkt zudem die leistungsbezogene Homogenität der Lernenden mit steigendem Jahrgang. Denn immerhin 11,2 Prozent der Gymnasialschülerinnen und -schüler wechseln zwischen dem 7. und dem 9./10. Schuljahr in einen anderen Bildungsgang der Sekundarstufe. Diese Prozesse der Abwärtsmobilität führen insgesamt dazu, dass die Leistungsmittelwerte im Gymnasium angehoben und die Streuung reduziert werden (Baumert et al. 2003). Auch psychosoziale Merkmale, wie Selbstwirksamkeit, fallen im Mittel an Gymnasien höher aus als an anderen Schularten (Köller 2007; Prenzel et al. 2013). Insgesamt stellen die Lernenden am Gymnasium damit eine unter kognitiven, psychosozialen und sozialen Gesichtspunkten positiv ausgewählte Schülerschaft dar.

Das Gymnasium bietet die größte Bandbreite an Anschluss- und Abschlussmöglichkeiten im traditionellen Bildungssystem, da das Abitur den Weg zum Studium öffnet und zugleich die aussichtsreichsten nichtakademischen Berufe erschließt (Liebau et al. 1997). Zudem erwerben die Lernenden an Gymnasien nach dem erfolgreichen Abschluss der 9. oder 10. Klasse (meist automatisch)

weitere Bildungsabschlüsse, wie den Hauptschulabschluss oder den Mittleren Abschluss (KMK 2017a). Je nach Bundesland sind hier bestimmte Notenkonsstellationen maßgebend oder in bestimmten Fällen Ergänzungsprüfungen notwendig. Trotz der vielfältigen Möglichkeiten verließen im Jahr 2014 86 % der Lernenden das Gymnasium mit dem Abitur. Lediglich 2 % bzw. 11 % beenden das Gymnasium mit dem Hauptschulabschluss bzw. dem Mittleren Abschluss (Autorengruppe Bildungsberichterstattung 2016). Mit der steigenden gymnasialen Beteiligung steigen auch die Anteile von Lernenden im Berufsbildungssystem. Der größte Anteil der Lernenden nimmt jedoch im Anschluss an das Abitur ein Studium auf (58 % im Jahr 2014). Damit bleibt das Gymnasium nicht nur in seinen historisch verankerten Wertorientierungen, sondern auch in der aktuellen Ausgestaltung die Schulart, welche vorrangig auf das Abitur und die universitäre Bildung vorbereitet.

2.2.4 Bildungsgangübergreifende Schulen

Bildungsgangübergreifende Schulen bieten im Gegensatz zum traditionellen dreigliedrigen Bildungssystem mehrere Bildungsgänge an einer Schule an. Die Organisationsformen unterscheiden sich dabei in der Anzahl angebotener Bildungsgänge (zwei vs. drei) sowie in der Durchlässigkeit zwischen den Bildungsgängen (integrierte vs. kooperative/additive Formen). Integrierte Einrichtungen sind Schulen, bei denen die Lernenden in Jahrgangsklassen gemeinsam unterrichtet werden. Der Unterricht in den Hauptfächern findet jedoch in der Regel in Kursen mit unterschiedlichen Niveaustufen statt. Kooperative bzw. additive Schulen bieten parallel mehrere Bildungsgänge an einer Schule an. Die Lernenden bzw. die Schulklassen können einem bestimmten Bildungsgang zugeordnet werden.

In den Schulstatistiken werden neben den traditionellen Schularten – Hauptschule, Realschule und Gymnasium – zwei weitere Schularten geführt und durch die Anzahl der angebotenen Bildungsgänge voneinander abgegrenzt. Einrichtungen der Sekundarstufe I werden als Schulen mit mehreren Bildungsgängen bezeichnet, wenn sie nur die Bildungsgänge von Haupt- und Realschulen anbieten. Alle Schulen mit integriertem Unterricht, welche drei Bildungsgänge anbieten, fallen unter die Kategorie der integrierten Gesamtschulen. Additive bzw. kooperative Schulen, welche drei Bildungsgänge anbieten (meist kooperative bzw. additive Gesamtschulen genannt) werden in der Schulstatistik

nicht gesondert ausgewiesen, sondern sind bei den entsprechenden Schularten – Hauptschule, Realschule und Gymnasium – enthalten (KMK 2017b). In der historischen Herleitung der Schularten wurde jedoch deutlich herausgestellt, dass sich das traditionelle dreigliedrige Schulsystem zeitlich vorgelagert zu den bildungsgangübergreifenden Schulen entwickelt hat. Zudem sind das traditionelle Bildungssystem und die neueren bildungsgangübergreifenden Schularten mit je eigenen politischen und gesellschaftlichen Entwicklungen und Ideologien verbunden. Diese Arbeit schließt bei der Kategorisierung der Schularten deshalb nicht an das System der Bildungsstatistiken an, sondern übernimmt die Kategorisierung, wie sie vom deutschen PISA-Konsortium realisiert wird. Hier werden inhaltliche Unterscheidungskriterien stärker berücksichtigt: Hauptschulen, Realschulen und Gymnasien sind hier Schulen, welche den einen entsprechenden Bildungsgang anbieten. *Integrierte Gesamtschulen* sind, übereinstimmend mit den Bildungsstatistiken, Schulen mit integriertem Stufenaufbau, welche drei Bildungsgänge anbieten. Alle übrigen Schulen mit zwei oder drei Bildungsgängen werden vom deutschen PISA-Konsortium unter die Kategorie der *Schulen mit mehreren Bildungsgängen* gefasst. Diese Schulen mit mehreren Bildungsgängen lassen nur wenig von der ursprünglichen Idee der Gesamtschule erkennen (Köller 2008), vielmehr sind hier pragmatische Gründe richtungsweisend. Eine Zuordnung der verschiedenen Organisationsformen zu den Schularten, wie sie vom deutschen PISA-Konsortium 2012 umgesetzt wurde, findet sich im Online-Anhang 12.2.

Im Folgenden werden die bildungsgangübergreifenden Schularten – integrierte Gesamtschule und Schulen mit mehreren Bildungsgängen – getrennt dargestellt. Bei der Interpretation der aufgeführten Statistiken müssen Unschärfen aufgrund der abweichend definierten Schularten bei den Bildungsstatistiken toleriert werden. Diese sollten die Kernaussage jedoch nicht belasten. Tabelle 1 veranschaulicht beispielhaft am Anteil der Lernenden in 9. Klassen die Abweichungen zwischen den Bildungsstatistiken und den PISA-Daten.

Tabelle 1. Verteilung der Lernenden in 9. Klassen auf die Schularten in Prozent.

	HS	RS	GY	IGS	MBG
Bildungsstatistik	18,36	26,02	36,43	10,17	8,59
PISA	13,34	22,79	35,90	9,23	16,27

Gegenüberstellung der Bildungsstatistiken zum Schuljahr 2011/2012 und den PISA-2012-Daten.

2.2.4.1 Integrierte Gesamtschule

Die integrierte Gesamtschule kann durch die Ablösung getrennter Bildungsgänge zugunsten gemeinsamen Unterrichts und Kursen mit verschiedenen Niveaustufen als Idealtypus der geforderten Reformen in den 1960er und 1970er Jahren gesehen werden. Damit ist diese Schulart vermutlich am stärksten geprägt von ideologischen, pädagogischen und bildungspolitischen Diskussionen. Trotzdem hatten Gesamtschulen lange quantitativ nur eine geringe Bedeutung (Köller 2008). In den letzten zehn Jahren ist jedoch ein deutlicher Anstieg der Besuchsquoten von 8,9 % im Schuljahr 2004/2005 hin zu 15,8 % im Schuljahr 2014/2015 zu verzeichnen (Malecki 2016). Grund hierfür sind unter anderem Umstrukturierungen von traditionellen Schularten in integrierte Gesamtschulen – teilweise unter neuen Begriffen wie der Gemeinschaftsschule, der integrierten Sekundarschule oder der Stadtteilschule (Expertengruppe zur Evaluation der Bremer Schulreform 2018; Maaz et al. 2013; Expertengruppe zur Evaluation der Bremer Schulreform 2018; Malecki 2016; Preuss-Lausitz 2008). Zudem sind integrierte Gesamtschulen, vor allem im Zuge der Bildungsexpansion, häufig die Wahl für Lernende, deren Bildungswege am Ende der Grundschulzeit noch nicht abzusehen sind (Köller 2008). Der Übertritt in die Gesamtschule erfolgt in der Regel im Anschluss an die Grundschule. Mit steigender Durchlässigkeit sind aber auch Wechsel von anderen Schularten der Sekundarstufe I möglich (KMK 2017a).

Programmatisch verfolgt die integrierte Gesamtschule eine ganze Bandbreite an Zielen. Die Forderungen aus den 1970er Jahren finden sich auch heute noch in den Schulprogrammen bzw. zusammenfassend in den Zielen der „Gemeinnützigen Gesellschaft Gesamtschulen“ (GGG). Pädagogisch sollen in „eine[r] Schule für alle“ die Lernenden optimal und nach individuellen Interessen und Fähigkeiten gefördert werden (Gemeinnützige Gesellschaft Gesamtschule 2015; Lang-Wojtasik et al. 2016). Zudem ist eine berufspraktische Vorbereitung ebenso Gegenstand wie eine breite Allgemeinbildung. Realisiert wird dies durch ein breites Angebot an Wahlpflicht- und Wahlkursen (Köller 2008). Neben den kognitiven Fähigkeiten sind erzieherische Aspekte hin zu reflektiertem Verhalten sowie ein demokratieeftendendes Umfeld erklärte Ziele (Gaude 1996; Gemeinnützige Gesellschaft Gesamtschule 2015; Heymann 2013). Sozial- und bildungspolitisch gilt es zudem, ungleiche Bildungschancen abzubauen und die Zahl gehobener Bildungsabschlüsse zu erhöhen (Gaude 1996; Lang-Wojtasik et al. 2016). Damit verbunden ist das Ziel, eine frühe Weichenstellung

einzudämmen. Curricular müssen integrierte Gesamtschulen daher einen hohen Grad an Abstimmung in den vermittelten Inhalten zwischen den leistungsdifferenzierten Kursen gewähren (Köller 2008).

Besucht werden integrierte Gesamtschulen von einer heterogenen Schülerschaft. Die leistungsbezogene Zusammensetzung an integrierten Gesamtschulen ist ähnlich heterogen wie die an Realschulen und Hauptschulen. In der mittleren Leistung schneiden integrierte Gesamtschulen im deutschlandweiten Vergleich jedoch besser ab als Hauptschulen und schlechter als Gymnasien oder Realschulen. Auffällig ist, dass auch die Leistungsspitze an integrierten Gesamtschulen nur die durchschnittlichen Leistungen der Lernenden an Gymnasien erreichen (Prenzel et al. 2013). Ähnliche Ergebnisse konnten durch die Evaluation der Schulreformen in Bremen und Berlin festgestellt werden (Expertengruppe zur Evaluation der Bremer Schulreform 2018; Maaz et al. 2013, 2013). Es ist daher zu vermuten, dass die leistungsstärksten Lernenden eher den Besuch eines Gymnasiums anstreben. Zudem scheinen integrierte Gesamtschulen für Lernende mit hohem sozialem Status weniger attraktiv zu sein, denn es wechseln deutlich mehr Lernende mit niedrigem und mittlerem sozialem Status nach der Grundschule an eine integrierte Gesamtschule (Autorengruppe Bildungsberichterstattung 2016; Expertengruppe zur Evaluation der Bremer Schulreform 2018; Maaz et al. 2013).

Die Bildungswege der Lernenden sind an integrierten Gesamtschulen offener als im traditionellen Bildungssystem. Dies zeigt sich auch an der hohen Mobilität der Lernenden in den leistungsdifferenzierten Kursen bis zur neunten Klasse (Köller 2008). Der größte Teil der Lernenden verlässt die integrierte Gesamtschule mit dem Mittleren Abschluss: 48,83 % im Schuljahr 2015/2016 (siehe Tabelle 2). Der Hauptschulabschluss sowie die Allgemeine Hochschulreife werden in etwa gleichen Anteilen erworben (21,96 % bzw. 23,97 %). Etwa 1 % der Lernenden mit Mittlerem Abschluss setzen ihre Bildungslaufbahn an einem traditionellen Gymnasium fort (Malecki 2016). Die Anteile sind jedoch stark von der Angebotsstruktur abhängig, weshalb aus diesen Zahlen keine Rückschlüsse auf die Erfolgsquote der integrierten Gesamtschulen gezogen werden können. Unterschiede in den Abschlüssen zeigen sich bspw. zwischen Schulen mit eigener Oberstufe und ohne eigene Oberstufe (Expertengruppe zur Evaluation der Bremer Schulreform 2018; Maaz et al. 2013) und zwischen den Bundesländern (Malecki 2016). Deutlich wird jedoch, dass die Varianz in den Ab-

schlüssen an integrierten Gesamtschulen höher ist als an den Schularten des traditionellen Bildungssystems.

Tabelle 2. Schulabgänger aus integrierten Gesamtschulen in Deutschland nach Abschlussart im Schuljahr 2015/2016.

Abgänger integrierter Gesamtschule nach Abschluss						
	Ohne Abschluss	Hauptschulabschluss	Mittlerer Abschluss	Fachhochschulreife	Allgemeine Hochschulreife	Gesamt
N	6161	25851	57482	0	28216	117710
%	5,23	21,96	48,83	0,00	23,97	100

Quelle: Statistisches Bundesamt GENESIS-Online Datenbank (2018): Absolventen/Abgänger: Bundesländer, Schuljahr, Geschlecht, Schulabschlüsse, Schulart. Eigene Berechnung für Gesamtdeutschland.

2.2.4.2 Schulen mit mehreren Bildungsgängen

In der Kategorie „Schulen mit mehreren Bildungsgängen“ findet sich vermutlich die heterogenste Gruppe von Organisationsformen der Sekundarstufe I. Auf der einen Seite versuchen diese Schulen, ähnlich den integrierten Gesamtschulen, das dreigliedrige Bildungssystem zu überwinden und stärker auf individuelle Interessen und Fähigkeiten der Lernenden einzugehen. Auf der anderen Seite sind sie jedoch weniger strikt in der Umsetzung. Denn es werden entweder nur zwei der drei Bildungsgänge angeboten, wodurch das Gymnasium hier seine exklusive Stellung behält (Preuss-Lausitz 2008), oder die drei Bildungsgänge werden an einer Schule angeboten, sind aber organisatorisch und unterrichtlich getrennt (Köller 1996). Beides verringert die Flexibilität in den Bildungswegen sowie die soziale Integration und Chancengerechtigkeit. Trotzdem werden Schulen mit mehreren Bildungsgängen von Eltern wie auch von der Bildungspolitik als Chance gesehen. Dies zeigt sich auf der einen Seite im Schulwahlverhalten der Eltern, welche sich von den Hauptschulen abwenden. Schulen mit mehreren Bildungsgängen verzeichnen hingegen steigende Schülerzahlen mit einem Anstieg von knapp 4 % innerhalb der letzten zehn Jahre auf 11,4 % im Schuljahr 2014/2015 (Malecki 2016). Auf der anderen Seite wächst auch das Angebot an Schulen mit mehreren Bildungsgängen, da bei derzeit sinkenden Schülerzahlen vor allem in strukturschwachen Gebieten differenzierte Bildungsangebote nur schwer aufrechterhalten werden können (Baumert et al. 2017; Hurrelmann 2013).

Die unterrichtliche Organisation an Schulen mit mehreren Bildungsgängen scheint überwiegend pragmatisch statt ideologisch geprägt. Die Ideen des dreigliedrigen Bildungssystems sowie der Gesamtschuldebatte finden sich in abgeschwächter und teils durchmischter Form wieder. Schulen mit zwei Bildungsgängen werden von Seiten der Ministerien mit einem starken Bezug zur beruflichen Vorbereitung definiert. Stärker als im traditionellen Schulsystem stehen an Schulen mit mehreren Bildungsgängen jedoch die Interessen und Fähigkeiten des einzelnen Lernenden im Vordergrund. Beispielhaft sind hier die Definitionen der Mittelstufenschule in Hessen sowie der Regionalschule in Mecklenburg-Vorpommern aufgeführt. Für die weiteren Organisationsformen der Schulen mit mehreren Bildungsgängen sei auf die Dokumentation der jeweiligen Länder verwiesen.

„Individuelle Förderung in kleineren Lerngruppen, praxis- und handlungsorientierter Unterricht auf Basis einer gesonderten Stundentafel sowie die Einbindung der beruflichen Bildung sind die kennzeichnenden Elemente der pädagogischen Arbeit in der Mittelstufenschule. Berufsorientierung und Förderung der Ausbildungsreife sind weitere wesentliche Bestandteile des gesamten Unterrichts.“ (Hessisches Kultusministerium 2018)

„Die Regionale Schule umfasst die Klassen 5 bis 10. Sie bereitet die Schülerinnen und Schüler auf eine Berufsausbildung vor. Gleichzeitig ermöglicht sie aber auch den Zugang zu weiterführenden Bildungsgängen. [...] Wahlpflichtunterricht ab der Klasse 7 dient der Förderung der besonderen Interessen, Neigungen und Begabungen der Schülerinnen und Schüler. Gewählt werden kann zwischen der zweiten Fremdsprache und einem von der Schule angebotenen Kurs aus den Bereichen Arbeit-Wirtschaft-Technik und Informatik, Naturwissenschaften, Gesellschaftswissenschaften, Künstlerisch-musisches Aufgabenfeld und Sport.“ (Bildungsserver Mecklenburg-Vorpommern 2018)

Besucht werden Schulen mit mehreren Bildungsgängen von einer ähnlichen heterogenen Schülerschaft wie sie an integrierten Gesamtschulen zu finden ist. Die mittlere Leistung ist jedoch geringfügig schlechter als an integrierten Gesamtschulen, was auf den teils fehlenden gymnasialen Bildungsgang zurückgeführt werden kann. Auch an diesen Schulen fehlen besonders leistungsstarke Lernende. Ähnlich wie an den integrierten Gesamtschulen können die leis-

tungsstärksten Lernenden der Schulen mit mehreren Bildungsgängen nur an die durchschnittlichen Leistungen der Lernenden an Gymnasien heranreichen (Prenzel et al. 2013). Auffällig ist die geringe Wahl dieser Schulart durch Lernende mit Migrationshintergrund. Im Schuljahr 2014/2015 besuchen 7,2 % der Lernenden mit Migrationshintergrund eine Schule mit mehreren Bildungsgängen. Damit liegt die Besuchsquote noch unter der für das Gymnasium (16,2 %). Hauptschulen (30,6 %), Realschulen (21,1 %) und integrierte Gesamtschulen (20,1 %) verzeichnen deutlich größeren Zulauf von Lernenden mit Migrationshintergrund. Der geringe Zulauf an Schulen mit mehreren Bildungsgängen kann vermutlich auf die Lage der Schulen mit zwei Bildungsgängen zurückgeführt werden. Thüringen, Sachsen, Sachsen-Anhalt, Brandenburg und Mecklenburg-Vorpommern stellen den größten Anteil an Lernenden in Schulen mit zwei Bildungsgängen (Tabelle siehe Online-Anhang 12.3). In diesen Bundesländern wohnen weniger Menschen mit Migrationshintergrund (6,33 %) als im gesamtdeutschen Durchschnitt (Statistisches Bundesamt 2017b).

Die Bildungswege der Lernenden an Schulen mit mehreren Bildungsgängen unterscheiden sich nach dem gewählten bzw. absolvierten Bildungsgang. An allen Schulen dieser Schulart kann der Hauptschulabschluss sowie der Mittlere Abschluss bzw. die Mittlere Reife erworben werden. Nur ein kleiner Teil bietet zudem einen gymnasialen Bildungsgang an, welcher an einer Schule mit Sekundarstufe II fortgeführt werden kann. Bei den Schulen, welche zwei Bildungsgänge anbieten, verlassen 23,44 % der Lernenden die Schule mit Hauptschulabschluss. 6,81 % gehen ohne Abschluss von der Schule ab. Der größte Anteil (69,76 %) verlässt die Schule mit dem Mittleren Abschluss (siehe Tabelle 3). Der inhaltliche Schwerpunkt in der Berufsvorbereitung, vor allem von Schulen mit zwei Bildungsgängen, lässt vermuten, dass ein Großteil der Lernenden im Anschluss an die Schule in die Berufsausbildung mündet. Aus den Statistiken lassen sich die Berufswege der Lernenden, welche Schulen mit mehreren Bildungsgängen besucht haben, jedoch nur unzureichend nachvollziehen.

Tabelle 3. Schulabgänger aus Schulen mit zwei Bildungsgängen in Deutschland nach Abschlussart im Schuljahr 2015/2016.

Abgänger aus Schulen mit zwei Bildungsgängen nach Abschluss						
	Ohne Abschluss	Hauptschulabschluss	Mittlerer Abschluss	Fachhochschulreife	Allgemeine Hochschulreife	Gesamt
N	5083	17502	52097	0	0	74682
%	6,81	23,44	69,76	0	0	100,01

Quelle: Statistisches Bundesamt GENESIS-Online Datenbank (2018): Absolventen/Abgänger: Bundesländer, Schuljahr, Geschlecht, Schulabschlüsse, Schulart. Eigene Berechnung für Gesamtdeutschland.

2.2.5 Zusammenfassende Gegenüberstellung der Schularten der Sekundarstufe I

Die obigen Beschreibungen geben ein Gesamtbild für jede Schulart. Um die wesentlichen Unterschiede zwischen den Schularten in den Leitbildern, den Bildungszielen und -inhalten, der Schülerschaft sowie den Anschlussperspektiven deutlicher hervorzuheben, stellt Abbildung 2 die Kernaussagen der Schulartbeschreibungen einander gegenüber. Hier wird deutlich, dass die Schularten durchaus Überschneidungen in einzelnen Bereichen aufweisen, bspw. die Leistungszusammensetzung der Schülerschaft an integrierten Gesamtschulen und Schulen mit mehreren Bildungsgängen oder die Anschlussperspektive nach dem Besuch einer Realschule, einer integrierten Gesamtschule oder einer Schule mit mehreren Bildungsgängen. Trotzdem ergeben sich im Gesamtprofil deutliche Unterschiede zwischen allen Schularten. Welche Konsequenzen sich hieraus für den Unterricht ergeben, wird im nächsten Kapitel näher beschrieben.

Abbildung 2. Die Schularten der Sekundarstufe I.

	Hauptschule	Realschule	Gymnasium	Integrierte Gesamtschule	Schulen mit mehreren Bildungsgängen
Leitbild	<ul style="list-style-type: none"> •Schulart mit geringstem Bildungsanspruch •Bedürfnisse des Individuums im Vordergrund •Praktische Vorbereitung auf handwerkliche und hauswirtschaftliche Berufe 	<ul style="list-style-type: none"> •Vorbereitung auf Berufe in Wirtschaft, Handel und Technik 	<ul style="list-style-type: none"> •Elitäre Bildungsanstalt •Ermöglicht die besten Anschlussperspektiven 	<ul style="list-style-type: none"> •Geringere Weichenstellung •Soziale Integration und Chancengleichheit •Förderung individueller Interessen und Fähigkeiten 	<ul style="list-style-type: none"> •Dreigliedriges Bildungssystem überwinden, aber exklusive Stellung der gymnasialen Bildung erhalten •Förderung individueller Interessen und Fähigkeiten •Pragmatische Reaktion auf bildungspolitische, soziale und demographische Hürden
Bildungsziele & Bildungsinhalte	<ul style="list-style-type: none"> •Berufspraktische Vorbereitung in eingeschränkten Berufsfeldern •Grundlegende Allgemeinbildung mit hoher Praxisorientierung 	<ul style="list-style-type: none"> •Allgemeinbildung mit Bezug zum praktischen Leben •Problemorientierter Unterricht •Berufsfeldbezogene Schwerpunkte •Erweitertes Angebot in Fremdsprachen und Naturwissenschaften (im Vergleich zur Hauptschule) 	<ul style="list-style-type: none"> •Vertiefende Allgemeinbildung •Vorbereitung auf Hochschulbildung •Fachlicher Zugang und Reflexionskultur •Kulturelle Bildung 	<ul style="list-style-type: none"> •Breite Allgemeinbildung •Berufsvorbereitung •Erzieherisches Wirken •Individuelle Förderung •Breites Kursangebot •Höherer Grad an Abstimmung zwischen leistungsdifferenzierten Angeboten 	<ul style="list-style-type: none"> •Breite Allgemeinbildung •Berufsvorbereitung •Individuelle Förderung
Selektion und Segregation	<ul style="list-style-type: none"> •Hoher Anteil an Lernenden aus bildungsfernen Elternhäusern •Hoher Anteil an Lernenden mit Migrationshintergrund •Niedrige durchschnittliche Leistung 	<ul style="list-style-type: none"> •Offen gegenüber allen sozialen Gruppen •Sinkender Anteil Lernender mit hohem sozialen Status •Mittlere durchschnittliche Leistung 	<ul style="list-style-type: none"> •Größter Anteil an Lernenden mit hohem sozialen Status •Hohe mittlere Leistung 	<ul style="list-style-type: none"> •Heterogene Leistungszusammensetzung bei fehlender Leistungsspitze •Größter Anteil an Lernenden mit mittlerem und niedrigem sozialen Status 	<ul style="list-style-type: none"> •Heterogene Leistungszusammensetzung bei fehlender Leistungsspitze •Größter Anteil an Lernenden mit mittlerem und niedrigem sozialen Status •Geringer Anteil an Lernenden mit Migrationshintergrund
Anschlussperspektive	<ul style="list-style-type: none"> •Berufliche Bildung überwiegend in Handwerk und Hauswirtschaft •Übergangssystem •Wechsel auf eine andere Schule der Sekundarstufe mit steigender Tendenz 	<ul style="list-style-type: none"> •Berufliche Bildung vorwiegend in anspruchsvollen Berufen im öffentlichen Dienst oder in Industrie und Handel •Wechsel in die Sekundarstufe II mit steigender Tendenz 	<ul style="list-style-type: none"> •Studium •Berufliche Bildung in anspruchsvollen Berufen im öffentlichen Dienst oder in Industrie und Handel 	<ul style="list-style-type: none"> •Hohe Mobilität in leistungsdifferenzierten Kursen bis zur 9. Klasse und große Varianz in den Abschlüssen •Berufliche Bildung •Wechsel in die Sekundarstufe II mit steigender Tendenz 	<ul style="list-style-type: none"> •Berufliche Bildung •Wechsel in die Sekundarstufe II mit steigender Tendenz

3 Schulartspezifische Ausgangsbedingungen für den Unterricht

Ziel des vorangegangenen Kapitels war es, das Bildungssystem der Sekundarstufe I in seiner historischen Entwicklung und in seiner aktuellen Ausgestaltung zu beschreiben und die Schularten zu kontrastieren. Es wurde deutlich, dass sich die Schularten in wesentlichen Merkmalen – den Leitbildern, den Bildungszielen und -inhalten, der Selektion sowie in den Anschlussperspektiven – voneinander unterscheiden. In diesem Kapitel gilt es nun, von der Ebene des Bildungssystems (Kapitel 2) überzuleiten auf das eigentliche Kerngeschäft von Schule, den Unterricht.

Nach Fend (2009) können vier gesellschaftliche Funktionen von Schule unterschieden werden: kulturelle Reproduktion, Qualifikation, Allokation und Integration. Um dem Anspruch dieser gesellschaftlichen Funktionen von Schule gerecht zu werden, zielt Beschulung laut Fend (2009) auf die langfristig angelegte Entwicklung von Kindern und Jugendlichen in verschiedenen Dimensionen. Aus den gesellschaftlichen Funktionen ergeben sich demnach individuelle Funktionen von Schule für die Lernenden. Durch die Qualifikationsfunktion der Schule beispielsweise werden den Lernenden Fähigkeiten und Fertigkeiten vermittelt, die grundlegend für die spätere Berufsfähigkeit sind. Um die Ziele des Bildungswesens umzusetzen, müssen die gesellschaftlichen und individuellen Funktionen ineinandergreifen (Fend 2009). In seiner Neuen Theorie von Schule kommt Fend zu dem Schluss, dass die entscheidende Verbindung zwischen gesellschaftlichen und individuellen Funktionen in der pädagogischen Arbeit liegt, als „Menschenbildung“ im Sinne der Förderung von Wissen, Kompetenzen, psychischen Ressourcen und Werten“ (Fend 2009, S. 169). Das zentrale Element hierfür ist der Unterricht.

Zur Beschreibung von Unterricht wird in der empirischen Forschung zur Effektivität von Unterricht häufig auf Angebots-Nutzungs-Modelle zurückgegriffen, welche betonen, dass das Unterrichtsangebot der Lehrperson und die Nutzung durch die Lernenden von verschiedenen Kontextfaktoren abhängig sind. „In diesem Sinne werden Unterricht und das Lehren bzw. Gestalten von Lernumgebungen als eine Angebotsstruktur aufgefasst, die von den Lernenden für sich genutzt werden will (soll). Die Art und Weise, wie Lernende die Angebotsstruktur nutzen, hängt von ihren individuellen Voraussetzungen und den Bil-

dungskontexten ab, in denen sie sich bewegen“ (Seidel 2014, S. 857). Angebots-Nutzungs-Modelle bilden häufig den theoretischen Hintergrund, wenn die Wirkung von Unterrichtsmerkmalen auf kognitive und affektiv-emotionale Merkmale der Lernenden untersucht wird. Für Untersuchungen, welche Zusammenhänge zwischen den Komponenten des Unterrichts sowie daraus resultierende Unterschiede im Unterricht fokussieren, scheinen diese Modelle hingegen weniger geeignet. Zum einen vernachlässigen Angebots-Nutzungs-Modelle häufig den Gegenstand als Komponente im interaktiven Prozess des Unterrichts. Zum anderen mangelt es an konkreten Überlegungen und empirischen Untersuchungen, wie einzelne Kontextmerkmale mit den Unterrichtsmerkmalen zusammenhängen.

In der Didaktik angesiedelte Unterrichtstheorien hingegen versuchen explizit zu fassen, wie einzelne Faktoren zusammenhängen. Das Didaktische Dreieck (Abbildung 3) fasst als theorieübergreifendes Modell die „Fundamentalstruktur der Unterrichtssituation“ (Sünkel 1996, S. 11) als ein relationales und interdependentes Geschehen zwischen dem Gegenstand (G), der Lehrperson (L) und dem/der Lernenden bzw. – auf den Klassenkontext bezogen – der Schülergruppe (S; Meseth und Proske 2011; Prange 2011). Die Interaktionen und die personalen Beziehungen zwischen Lehrenden und der Schülergruppe (L-S) sind dabei auf einen Gegenstand gerichtet (G). Die Lehrperson bereitet die Gegenstände (L-G) im Hinblick auf die Fähigkeiten der Schülergruppe (S) auf. Diese setzt sich wiederum mit dem Gegenstand (S-G) unter der Instruktion des Lehrenden (L) auseinander.

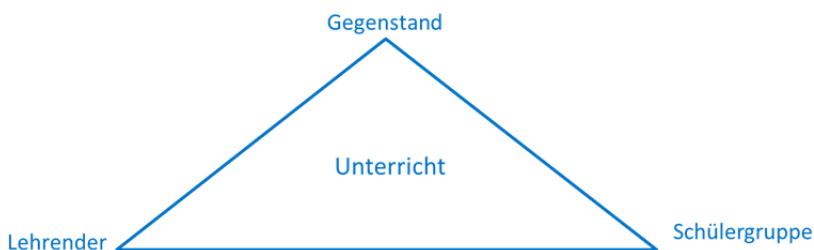


Abbildung 3. Didaktisches Dreieck.

Das Didaktische Dreieck stellt den Unterricht in reduzierter Komplexität dar, betont aber die wesentlichen Komponenten in ihrer Interdependenz. Damit ist es geeignet, „die Besonderheit der unterrichtlichen Situation zu kennzeichnen“

(Prange 1983, S. 36). Zugleich wird betont, dass die Ausgestaltung des Unterrichts von der jeweiligen Ausgestaltung der drei Komponenten abhängt.

Allerdings betrachtet das Modell in seiner allgemeinen Darstellung das Handeln unabhängig von der Umwelt. Unterricht ist jedoch institutionell vorstrukturiert und vollzieht sich im Kontext sozialer Organisationen. Für eine vergleichende Beschreibung und Analyse des Unterrichts muss deshalb die kontextuelle, institutionelle und organisatorische Rahmung des Unterrichts sowie die soziale Dynamik der Klasse mitgedacht werden (Herzog 2010).

Um die interaktiven Prozesse des Unterrichts im Kontext der schulartspezifischen Bedingungen darzustellen, wird deshalb das Didaktische Dreieck um Kontextmerkmale des Bildungssystems erweitert. Hierfür werden die herausgearbeiteten Unterschiede zwischen den Schularten (siehe Kapitel 2), mit den Komponenten des Didaktischen Dreiecks in Bezug gebracht. Daraus wird ersichtlich, dass sich für den Unterricht in den verschiedenen Schularten unterschiedliche Ausgangsbedingungen für die Komponenten des Unterrichts und somit für die Unterrichtsprozesse selbst ergeben (Abbildung 4).

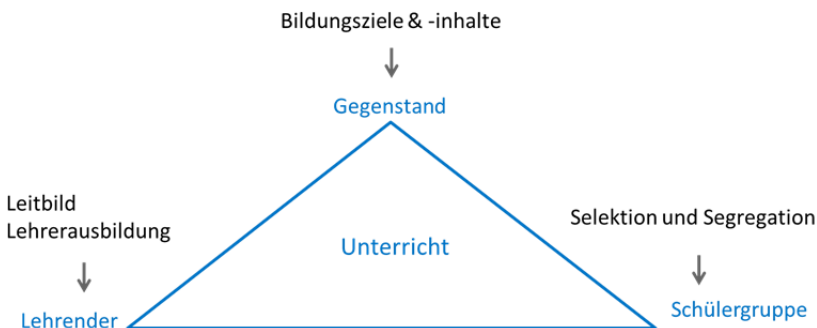


Abbildung 4. Didaktisches Dreieck mit schulartspezifischen Ausgangsbedingungen.

Ähnliche Einflüsse durch schulartspezifische Ausgangsbedingungen auf das schulische Lernen nehmen Baumert und Kollegen an, wenn sie von Differenziellen Lern- und Entwicklungsmilieus sprechen:

„Wenn wir von differenziellen Lern- und Entwicklungsmilieus sprechen ist damit gemeint, dass junge Menschen *unabhängig von und zusätzlich zu ih-*

ren unterschiedlichen persönlichen, intellektuellen, kulturellen, sozialen und ökonomischen Ressourcen je nach besuchter Schulform differenzielle Entwicklungschancen erhalten, die schulmilieubedingt sind und sowohl durch den Verteilungsprozess als auch durch die institutionellen Arbeits- und Lernbedingungen und die schulformspezifischen pädagogisch-didaktischen Traditionen erzeugt werden.“ (Baumert et al. 2006b, S. 98–99)

Empirisch stützt sich die Beschreibung und Analyse der Lern- und Entwicklungsmilieus vorwiegend auf die Daten aus den Studien BIJU (Bildungsverläufe und psychosoziale Entwicklung im Jugendalter), TIMSS 1999 (Trends in International Mathematics and Science Study), PISA (Programme for International Student Assessment) sowie die deutsche klassenbasierte Erweiterung zur PISA-Erhebung im Jahre 2003, genannt „PISA-E“. Zusätzlich zu dem Einfluss von institutionellen Effekten wie Lehrplan, Lern- und Unterrichtskultur sowie Effekten der Schülerkomposition werden auch die individuellen Voraussetzungen der Schülerinnen und Schüler als relevant für die Unterschiede in den Lernfortschritten gesehen (Baumert et al. 2006b; Baumert und Köller 1998; Stanat et al. 2010).

Baumert und Kollegen untersuchen aufbauend auf ihrer These vorwiegend Merkmale der Schülergruppe, um unterschiedliche Ausgangsbedingungen im schulischen Lernumfeld zu beschreiben (Baumert et al. 2006b; Baumert et al. 2009; Baumert et al. 2010; Baumert und Schümer 2001; Becker et al. 2012; Köller und Baumert 2001b; Kunter et al. 2005; Kunter et al. 2006). Weniger hergeleitet und untersucht werden die konkreten Prozesse auf Klassenebene sowie mögliche Zusammenhänge dieser mit den Entwicklungsmöglichkeiten der Lernenden. Zudem sind Überlegungen zum Lernumfeld bezogen auf die Unterrichtsinhalte und das Lehrerhandeln in den Arbeiten kaum berücksichtigt. Genau hier setzen die theoretischen Überlegungen dieser Arbeit an. Der folgende Abschnitt beschreibt einleitend das Verständnis von Unterricht, welches dieser Arbeit zugrunde liegt und beschreibt die zentralen Merkmale zur Analyse unterrichtlicher Prozesse (Abschnitt 3.1). Die anschließenden Abschnitte gehen näher auf die einzelnen Komponenten – Lehrende, Gegenstand und Schülergruppe – in ihrem Zusammenhang mit dem unterrichtlichen Lehren und Lernen ein und reflektieren potentielle Unterschiede zwischen den Schularten (Abschnitte 3.2 bis 3.4). In einem abschließenden Abschnitt wird der Forschungsstand empirischer Schulartunterschiede im Lehren und Lernen zusammengefasst (Abschnitt 3.5).

3.1 Unterricht und Unterrichtsprozesse

An einer befriedigenden Theorie des Unterrichts sowie einer einheitlichen Diskussion des Gegenstandes Unterricht scheint es bis heute zu mangeln. Dies zeigt sich zum einen in Gegenüberstellungen von Unterrichtsdefinitionen, wie es bspw. Lüders (2012) und Schramm (1975) vorgenommen haben. Beide finden übereinstimmend eine hohe sprachliche Heterogenität der Definitionen sowie unterschiedliche Schwerpunktsetzungen auf Lehrtätigkeiten (wie Vermittlung, Lehre, Einflussnahme, Führungshilfe oder Erziehung) oder Unterrichtsprozesse (wie Kommunikation, Lernen, Sozialisation, oder das Zusammenspiel von Lehren und Lernen). Zum anderen wird die Heterogenität der Perspektiven auf Unterricht und die Komplexität des Gegenstandes deutlich an Versuchen, die Unterrichtstheorien zu ordnen (Eickhorst 2011; Scholl 2011), Bausteine für eine zu entwickelnde Theorie des Unterrichts zu identifizieren (Breidenstein 2010), oder den Unterrichtsbegriff disziplinübergreifend zu diskutieren, wie im Streitgespräch zwischen Andreas Gruschka, Walter Herzog, Wolfgang Meseth, Matthias Proske und Sabine Reh geschehen (Meseth et al. 2011). Neben dem biografischen Alltagsverständnis von Unterricht prägen wissenschaftliche Diskurse aus der Allgemeinen Didaktik, der pädagogischen Psychologie sowie der Soziologie die Sicht auf das komplexe Phänomen Unterricht und thematisieren verschiedene Schwerpunkte, wie die Aufbereitung des Gegenstandes, das Lehrerhandeln oder die sozialen Interaktionen (Breidenstein 2010).

Diese Arbeit kann zu einer übergreifenden Theorie des Unterrichts nicht beitragen, wird aber in Kenntnis der verschiedenen Perspektiven und Komplexität des Phänomens die hier verwendete Sichtweise auf Unterricht für die weitere Argumentation abgrenzen. Unterricht wird hier als interaktiver Prozess zwischen Lernenden, Lehrenden und dem Gegenstand definiert im Kontext von Schule und den damit einhergehenden Regeln und Strukturen. Die ausgebildete Lehrperson plant und initiiert die Abfolge von Lehr- und Lernsituation. Die Lernenden einer Klasse gestalten mit ihren Voraussetzungen und ihrem Verhalten den Handlungsspielraum der Lehrperson (Lipowsky 2015; Terhart 1994). Unterricht als sozialer Prozess wie auch das verhandelte Wissen stellen daher eine Ko-Produktion der beteiligten Personen dar (Klieme 2006), wobei der Fokus dieser Arbeit auf dem Lehrerhandeln liegt, welches in Interaktion mit den Lernenden und dem Gegenstand vollzogen wird. Die Interaktionen des Unterrichts beziehen sich hier auf eine zeitlich und örtlich festgelegte Situation

in Form von Unterrichtsstunden und schließen intendierte wie nicht intendierte Interaktionen ein, bspw. den Umgang mit Störungen. Damit ist diese Sichtweise auf Unterricht breiter, als sie bspw. Gruschka (2013) fasst, wenn er von „Unterrichten“ spricht: „wohl wissend, dass im Unterricht manches andere kategorial bestimmende auftreten kann: etwa Beratung, Therapie, Unterhaltung oder Spiel, Arbeiten oder schiere Beschäftigung. [...] Wenn alles Unterrichten genannt wird, dann erübrigt sich eine Theoretisierung“ (Gruschka 2013, S. 50–51). Das Ziel von Unterricht wird im Aufbau von Wissen sowie im Erwerb von fachspezifischen und fächerübergreifenden Kompetenzen auf Seiten der Lernenden gesehen. Zudem soll Unterricht die Lernenden in unterschiedliche (fachliche) Modi des Denkens und Handelns einführen und Bildung im Sinne kultureller Sozialisation und mündiger Teilhabe ermöglichen (Klieme 2018; Lipowsky 2015; Terhart 1994).

Häufig eng verbunden mit den Zielen des Unterrichts ist die normative Setzung, was guter oder effektiver Unterricht ist (Klieme 2018), besonders in der Lehr-Lernforschung. Berliner (2016) unterscheidet hierbei die Kriterien für effektiven Unterricht von gutem Unterricht. „Effektiv“ bezieht er auf Leistungsziele, „guten Unterricht“ begründet er hingegen über erfüllte Verhaltenserwartungen bspw. bezogen auf den Umgang mit den Lernenden oder hinsichtlich bestimmter Methoden. In der Lehr-Lernforschung wird jedoch in den Analysen Effektivität als auch Qualität meist auf die Leistung(sentwicklung) bezogen, gelegentlich auf lernrelevante psychosoziale Merkmale wie Motivation, Interesse oder Selbstwirksamkeit (Lipowsky 2015). Auch in der Didaktik oder der qualitativen Bildungsforschung wird die Frage nach der Unterrichtsqualität gestellt, bspw. wenn Klafki (1976) mittels pädagogischer Handlungsforschung prüft, ob der Unterricht seinen Gegenstand für die Lernenden und die Lernenden für den Gegenstand erschließt, oder wenn Gruschka (2013) Unterrichtssequenzen anhand einer Graduierung aus misslingender Vermittlung über Einübung und Verwicklung zu gelingender Vermittlung beschreibt (Klieme 2018). In dieser Arbeit liegt der Fokus hingegen stärker auf der Beschreibung und Kontrastierung von Unterricht. Nichtsdestotrotz schließt die Beschreibung von Unterricht über Merkmale des Lehrerhandelns an Befunde der empirischen Qualitätsforschung an, denn es gilt Merkmale einzubeziehen, welche sich bewährt haben, unterschiedliche und unterscheidbare Facetten des Unterrichts zu identifizieren. So wird auch bei der Übersicht zum Forschungsstand die normative Wertung einiger Theorien nicht verborgen bleiben. Es sei jedoch angemerkt, dass Effektivität oder Qualität von Unterricht im Auge des Betrachters

liegen (Meyer 2017, S. 11–12). Dies zeigt sich an den sehr unterschiedlichen Forschungsansätzen, Unterrichtsqualität zu beschreiben, lässt sich aber auch erahnen, wenn Unterricht aus Sicht anderer Akteure betrachtet wird. Aus Sicht der Lernenden könnte die Bewertung von gutem Unterricht beispielsweise stärker auf den Unterhaltungswert, die Relevanzzuschreibung oder die soziale Eingebundenheit abzielen. Aus einer Gerechtigkeitsperspektive heraus ist Unterricht effektiv oder gut, wenn Leistungsentwicklung und Abschlüsse unabhängig von der sozialen Herkunft sind (Fend 1982).

In der Lehr-Lernforschung haben besonders Unterrichtsmethoden sowie generische Faktoren des Unterrichts Bedeutung für die Beschreibung und Analyse von Unterrichtsprozessen erlangt (Decristan et al. 2015). Unterrichtsmethoden sind klar abgrenzbar, standardisierbar und daher empirisch gut zu untersuchen (Klieme 2018). In der Tradition der schweizerischen Unterrichtsforschung werden diese als „Sichtstruktur“ des Unterrichts bezeichnet (Reusser und Pauli 2013, S. 317). Ein weiteres Unterrichtsmerkmal, welches teilweise den Unterrichtsmethoden zugeordnet wird (Decristan et al. 2015), soll hier gesondert aufgeführt werden: Das formative Assessment. Dieses ist eher als Instrument für die Lehr- und Lernplanung zu betrachten, greift aber hierfür auf leicht identifizierbare Methoden zurück. Generische Faktoren des Unterrichts (Raudenbusch 2008) gelten als Umsetzung von Inhalten und Methoden (Klieme 2018). Sie beschreiben die Tiefenstrukturen des Unterrichts (Reusser und Pauli 2013).

3.1.1 Unterrichtsmethoden

Unterrichtsmethoden als „Formen und Verfahren, in und mit denen sich Lehrer und Schüler die sie umgebende natürliche gesellschaftliche Wirklichkeit unter institutionellen Rahmenbedingungen aneignen“ (Meyer, 1988, S. 45), sind als Merkmale der Sichtstrukturen von Unterricht leicht beobachtbar und vielfältig (Kunter und Trautwein 2013). Umso schwieriger ist eine Ordnung oder Kategorisierung der Formen und Verfahren. Meyer (2002) schlägt deshalb ein Ordnungsraster vor, das die Methodenvielfalt auf drei Ebenen einordnet. Auf Makroebene finden sich „methodische Großformen“, wie etwa der Projektunterricht, welche sich über mehrere Unterrichtsstunden hinweg ziehen können. Auf Mesoebene ordnet er das „methodische Handeln“ an. Hiervon grenzt er auf Mikroebene Inszenierungstechniken ab, welche oft nur wenige Sekunden dauernde Lehr-Lern-Situationen bezeichnen (Meyer 2002, S. 111). Die folgende

Beschreibung von Unterrichtsmethoden fokussiert auf das methodische Handeln, welches die von Seiten der Lehrkraft intendierten Handlungsmuster adressiert und die Gestaltung längerer in sich geschlossener Unterrichtssequenzen einschließt (Meyer 2002, 2017).

Bei der Auswahl der eingesetzten Unterrichtsmethoden stehen der Lehrperson eine Vielzahl von Sozial- und Handlungsformen zur Verfügung, welche sich stärker am Lehrenden oder an den Lernenden orientieren können (Wiechmann und Wildhirt 2016, 18ff). Stärker lehrerzentrierte Unterrichtsmethoden gehen auf Grundannahmen der Planbarkeit des methodischen Unterrichtsgeschehens im Sinne Herbarts (1805/1919) zurück. Herbartianer blendeten die unmittelbare Erfahrungswelt der Lernenden bewusst aus, um planvoll und ohne die Zufälligkeiten des täglichen Lebens einen erziehenden Unterricht realisieren zu können (Wiechmann und Wildhirt 2016). Der Unterricht ist damit mit einer starken Lehrerlenkung verbunden. Die Lehrkraft steuert die Prozesse, regt die Lernenden an und kontrolliert diese. Sie vermittelt den Lehrstoff, indem sie diesen vorträgt, vorliest, erzählt oder demonstriert. Inhalte, Lerntempo und Ziele werden durch die Lehrperson bestimmt und sind für alle Lernenden gleich. Die Lernenden sind in dieser Unterrichtsform eher reaktiv, in dem sie Inhalte aufnehmen (können) und Anweisungen folgen. In dieser Form können komplexe Sachverhalte didaktisch aufbereitet und vermittelt sowie die Konzentration der Lernenden auf zentrale Inhalte gelenkt werden (Kunter und Trautwein 2013; Meyer 2002). Dadurch kann die Lehrperson schnell mit ihrem Stoff durchkommen. Aber ob der Stoff auch bei den Lernenden ankommt, ist nicht garantiert (Meyer 2002). Lehrerorientiertem Unterricht lassen sich Methoden des Lehrervortrages, des fragend entwickelnden Unterrichtsgesprächs oder der direkten Instruktion zuordnen (Kunter und Trautwein 2013).

Stärker schülerorientiert sind Methoden, welche Wiechmann und Wildhirt (2016) reformpädagogischen Ansätzen sowie dem Konstruktivismus zuschreiben. Reformpädagogische Ansätze betonen das Erfahrungslernen, weshalb Unterricht in geringerem Maße von der Lehrperson alleine geplant werden kann. Der Lehrkraft kommt hier vielmehr eine prozessmoderierende Funktion zu, wie beispielsweise bei Projektunterricht, Werkstattarbeit, Erkundungen oder Gruppenarbeit. Konstruktivistische Ansätze betonen Autonomie und Selbstbestimmung. Hier werden Unterrichtsmethoden der Portfolioarbeit und des autonom-kooperativen Arbeitens eingeordnet (Miller 2013). Damit stehen bei schülerorientierten Unterrichtsmethoden die Lernenden stärker im Fokus. Die unter-

richtsbezogenen Entscheidungen sollen in erster Linie an den Lernenden ausgerichtet sein, an seiner Erfahrungswelt, seinen Bedürfnissen, seiner Ausgangslage, seinen Interessen und Erwartungen. Ziel ist es, die Lernenden aus der Rolle des bloßen Adressaten herauszulösen und stattdessen zur selbstständigen Informationsbeschaffung und -verarbeitung anzuleiten (Wiater 2014).

Eine vieldiskutierte Frage ist, welche Unterrichtsmethoden zu favorisieren sind. Hierzu gibt es – vor allem im englischsprachigen Raum – zahlreiche Publikationen, welche lehrerzentrierte Methoden den schülerorientierten Methoden gegenüberstellen, meist mit der Intention, die eine oder die andere Orientierung als überlegen darzustellen (bspw. Derting und Ebert-May 2010; Elen et al. 2007; Secker und Lissitz 1999). Im deutschsprachigen Raum wurde die Frage meist theoretisch angegangen (Bohl 2000). Unter sozialkonstruktivistischer Perspektive, die annimmt, dass Wissen durch soziale Konstruktionsprozesse entsteht (Gegern 1995; Youniss 1982), werden vor allem kooperative Arbeitsformen favorisiert, die eine symmetrische Kommunikation ermöglichen sollen, wie sie meist in eher schülerorientiertem Unterricht realisiert wird. Demgegenüber haben die Lernenden im lehrerzentrierten Unterricht aufgrund der asymmetrischen Kommunikation nur eingeschränkt die Möglichkeit, Wissen aktiv zu konstruieren (Youniss 1982). Empirisch zeigt sich, dass in stark lehrerzentriertem Unterricht die verbale Interaktion durch kurze Schülerantworten auf wenig komplexe Lehrerfragen bestimmt ist (Reyer 2004). Andererseits lassen Lehrpersonen, welche Sachverhalte gern geschlossen selbst darlegen (wie bspw. im Frontalunterricht), Lernende häufiger diskutieren (Blömeke und Müller 2009), was durch gesteigerte Reflexion und Verarbeitungstiefe wiederum als lernförderlich zu sehen ist.

Bis heute konnte weder empirisch noch in der didaktischen Diskussion eine der Orientierungen überzeugen. Stattdessen überwiegt inzwischen die Überzeugung, dass jede Methode mit ihren Vor- und Nachteilen reflektiert und auf die Situation hin ausgewählt werden sollte. „Die Unterschiedlichkeit der Lernvoraussetzungen, -fähigkeiten und -stile der Schüler darf nicht durch die Monokultur der einen Methode nivelliert werden. Vielmehr sind methodische Variationen einzusetzen, die eine Anpassung des Unterrichts an die individuellen Unterschiede erlauben“ (Weinert 1997, S. 50). Das bedeutet, dass Lehrkräfte über ein breites Repertoire an verschiedenen Methoden verfügen sollten, die sie kompetent und flexibel anwenden – je nach aktueller Zielsetzung und gegenwärtigen Bedingungen in der Klasse; es wird auch von der „Orchestrierung“

diverser didaktischer Strategien und methodischer Grundformen gesprochen (Einsiedler 1997; Helmke 2007).

3.1.2 Formatives Assessment

Ein weiteres Unterrichtsmerkmal, welches auch als Unterrichtsmethode angesehen werden kann (Decristan et al. 2015), aber schwerlich zu obigen Methoden zuzuordnen ist, soll hier gesondert aufgeführt werden: das formative Assessment. Als gut beobachtbares Merkmal, welches „on the fly“, aber auch eingebettet in den Unterricht auftreten kann (Ayala et al. 2008), ist es den Sichtstrukturen des Unterrichts zuzuordnen. Formatives Assessment dient jedoch nicht direkt der Vermittlung oder Aneignung von Inhalten, sondern ist eher als übergreifende Maßnahme zur Lehr- und Lernplanung zu sehen. Zentral ist, wie Informationen über das Wissen und Können der Lernenden genutzt werden (Maier, 2010a). Der Begriff „formatives Assessment“ wurde maßgeblich von Black und Wiliam (1998a, 1998b) geprägt, welche dieses definieren: „as encompassing all those activities undertaken by teachers, and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged“ (Black und Wiliam 1998a, S. 7). Somit kann formatives Assessment als Methode gesehen werden, die hilft, das eigene Lernen oder die Unterrichtsgestaltung besser zu planen. Feedback spielt beim formativen Assessment eine Schlüsselrolle, um den Soll/Ist- Zustand abzugleichen und um Strategien abzuleiten, die geeignet sind um den Soll-Zustand zu erreichen.

Die lerntheoretische Grundlage von formativem Assessment hängt nach James (2006) von der konkreten Gestaltung ab. Assessment, das im Sinne einer Fertigkeitshierarchie aufeinander aufbauende Testitems verwendet, Leistung als entweder richtig oder falsch kategorisiert und primär personenbezogenes Lob rückmeldet, ist enger mit behavioristischen Lernansätzen verknüpft als Assessment, das authentische Aufgaben einsetzt und Gruppenlernen sowie Peer- und Selbstassessment nutzt. Letzteres setzt eher sozial-konstruktivistische Auffassungen von Lernen um (James 2006) und wird als überlegenere Form des formativen Assessments favorisiert (Shepard 2000). Durch formatives Assessment können Entwicklungs- bzw. Lernspielräume identifiziert und damit der Lernende in seinen Entwicklungs- und Lernprozessen unterstützt werden. Wiliam und Thompson (2008) beschreiben fünf Schlüsselmerkmale des forma-

tiven Assessments: (a) Lernziele und Erfolgskriterien klären, teilen und verstehen; (b) Lernstand durch Diskussionen, Fragen und Aufgaben erfassen; (c) lernförderliche Rückmeldung geben; (d) Lernende als Verantwortliche des eignen Lernens aktivieren; (e) Lernende als instruktionale Ressourcen füreinander aktivieren. Aus diesen Merkmalen wird der starke Fokus auf den Schüler bzw. die Schülerin deutlich. Formatives Assessment und Feedback werden in der Forschung oftmals eingeschränkt hinsichtlich des Lernstandes und der Leistungsentwicklung von der Lernenden diskutiert (Kingston und Nash 2011). Formatives Assessment ist jedoch zugleich für die Unterrichtsplanung nützlich, um den Unterricht an den Bedürfnissen der Lernenden auszurichten.

Entgegen der zuvor beschriebenen Unterrichtsmethoden kann für formatives Assessment ein Zusammenhang mit der Leistungsentwicklung angenommen werden. Das Review von Black und Wiliam (1998a) ist der wohl am häufigsten zitierte Artikel, um den positiven Effekt von formativem Assessment auf Leistung zu belegen. Weitere Erkenntnisse liefern Kingston und Nash (2011) mit ihrer Metaanalyse basierend auf aktuellen Studien (Brookhart et al. 2008; Rakoczy et al. 2008). Auch sie fanden einen kleinen bis mittleren positiven Effekt des formativen Assessments auf Leistung. Darüber hinaus können sie zeigen, dass die Wirksamkeit von formativem Assessment vom Schulfach sowie von der Assessmentart abhängt. An Untersuchungen, welche den Einfluss von formativem Assessment auf die inhaltliche und methodische Gestaltung des Unterrichts prüfen, scheint es hingegen zu mangeln. Stattdessen findet sich Forschung über das Zusammenspiel globaler Faktoren der Unterrichtsprozesse mit formativem Assessment (Decristan et al. 2015). Hier zeigt sich, dass eine allgemein hohe kognitive Aktivierung im Klassenraum und ein unterstützendes Klima die positive Wirkung von formativem Assessment auf das Lernen der Schülerinnen und Schüler erhöhen. Den generischen Faktoren des Unterrichts, welche den Unterrichts als Ganzes konstituieren und somit nicht nur einzelne Sequenzen betreffen, kommt damit eine besondere Bedeutung zu (Klieme 2018).

3.1.3 Generische Faktoren des Unterrichts

Mit den beschriebenen schüler- oder lehrerorientierten Methoden oder der Methode des formativen Assessment lassen sich Unterrichtsstunden in ihren verschiedenen Gestaltungsmöglichkeiten beschreiben. Allerdings sind dies nur

einzelne Elemente, die nicht den Unterricht als Ganzes konstituieren können (Klieme 2018). In der Lehr-Lernforschung werden deshalb zunehmend generische Faktoren des Unterrichts diskutiert, welche in ihrer Anlage methodenübergreifend auf Schulfächer, Klassenstufen und möglicherweise auch auf Länder und Kulturen anwendbar sein sollen (Praetorius et al. 2018).

Im deutschen Sprachraum hat sich der Ansatz von Klieme et al. (2001) positioniert, welcher die Dimensionen konstruktive Unterstützung⁸, kognitive Aktivierung und Klassenführung⁹ als Qualitätsdimensionen des Unterrichts benennt. Unterrichtstheoretisch lassen sich diese drei Dimensionen mit Diederich und Tenorths (1997) Konstitutionsbedingungen von Unterricht begründen: „Demnach haben Lehrer drei Möglichkeiten, die Lernbereitschaft sicherzustellen, ohne die der Unterricht nicht beginnen und weitergeführt werden kann. Diese ‚Konstitutionsbedingungen‘ des Unterrichts sind Disziplin, Leistung und Motivation“ (Diederich und Tenorth 1997, S. 98), wobei Disziplin und Aufmerksamkeit durch eine gut strukturierte Klassenführung hergestellt werden können, Motivation durch konstruktive Unterstützung und Leistung durch ein hohes Maß an kognitiver Aktivierung (Klieme 2018). Empirisch wurden die drei Dimensionen konstruktive Unterstützung, kognitive Aktivierung sowie Klassenführung anhand videobasierter Untersuchungen von Mathematikstunden hergeleitet. Verschiedene Untersuchungen zeigen jedoch die Übertragbarkeit dieser Dimensionen auf anderer Fächer (bspw. Rjosk et al. 2014). Die psychologischen Mechanismen, die den Auswirkungen der Dimensionen auf die Schülerergebnisse zugrunde liegen, sind über gut etablierte kognitive und motivationale Theorien konkretisierbar (Praetorius et al. 2018). Diese werden im Folgenden näher ausgeführt.

Die Dimension konstruktive Unterstützung bezieht sich auf die Selbstbestimmungstheorie von Deci und Ryan (1996), welche die Motivation für ein bestimmtes Verhalten auf die Befriedigung von drei Grundbedürfnissen zurückführen: das Bedürfnisse nach Kompetenz, nach sozialer Eingebundenheit und nach Autonomie. Ein hohes Ausmaß an konstruktiver Unterstützung wird angenommen, wenn die Lehrkraft sensitiv für die individuellen Bedürfnisse der Lernenden ist, unterstützend agiert und geringen Leistungsdruck aufbaut

8 In manchen Studien auch „Schülerorientierung“ oder „Teacher support“ genannt. Um eine Verwechslung mit schülerorientierten Unterrichtsmethoden zu vermeiden, wird hier der Begriff konstruktive Unterstützung verwendet.

9 In manchen Studien auch „Strukturiertheit“ genannt.

(Klieme 2018). Zu dem zeichnet sich konstruktive Unterstützung aus durch eine wertschätzende Schüler-Lehrer-Beziehung sowie durch konstruktives Feedback, individuelle Unterstützung und eine positive Fehlerkultur (Klieme et al. 2009; Klusmann et al. 2008). Unterricht, der die Grundbedürfnisse nach Deci und Ryan (1996) befriedigt, erhöht die intrinsische Motivation der Lernenden (Dietrich et al. 2015) und das fachspezifische Interesse (Fauth et al. 2014).

Klassenführung bezieht sich auf die Qualität der Unterrichtszeit durch Struktur und Klarheit. Hierzu zählen, klare Regeln für erwünschtes Schülerverhalten im Sinne von aktiver Beteiligung und Aufmerksamkeit einzuführen, dieses z. B. durch Routinen zu stützen, sowie Störungen rechtzeitig zu erkennen und präventiv zu vermeiden (Klieme 2018). Im Sinne von Kounin (1970) bezieht sich die Klassenführung damit nicht nur auf die Reaktion der Lehrkräfte auf Störungen, sondern auch auf den gesamten Unterricht, in dem das Auftreten von Störungen im Klassenraum durch effektive Nutzung der Zeit, klar definierte Regeln, stabile Routinen, Unterrichtsdisziplin und Überwachung gering gehalten werden sollte (Praetorius et al. 2018). Klassenführung ist in diesem Sinne auch in offenen Lernsituationen erforderlich, damit die nominelle Lernzeit besser genutzt werden kann (Kuger 2016)

Die letzte Dimension, die kognitive Aktivierung, fasst Unterrichtspraktiken zusammen, die das Denken auf höherer Ebene fördern und die Metakognition unterstützen, indem sie herausfordernde Aufgaben und Fragen anwenden oder das Vorwissen der Schüler aktivieren und erforschen (Fauth et al. 2014; Klieme et al. 2009). Kognitiv aktivierender Unterricht zeichnet sich vor allem durch anspruchsvolles Üben, herausfordernde Aufgaben und die Nutzung von Fehlern als Lerngelegenheiten aus und soll die Lernenden anregen, Informationen zu rekonstruieren, auszuarbeiten und zu integrieren, was zu einem flexibleren Wissen und einem tieferen Verständnis führt (Praetorius et al. 2018). Erfasst wird der Grad der kognitiven Aktivierung in der Unterrichtsforschung meist über das eingesetzte Aufgabenmaterial oder den Einsatz bestimmter kognitiv aktivierender Unterrichtspraktiken. Deshalb wird dieses Merkmal in manchen Studien auch als „Potential zur kognitiven Aktivierung“ bezeichnet, da die kognitive „Aktiviertheit“ der Lernenden nicht gemessen werden kann (Kunter und Trautwein 2013).

Zusammenfassend zeigt sich, dass alle drei Dimensionen wichtige Elemente für einen kognitiv und motivational förderlichen Unterricht sind und im Sinne Diederich und Tenorths (1997) als konstituierend für Unterricht betrachtet werden können. Trotzdem bleibt Unterricht anhand der aufgeführten Merkmale ein recht abstrakter Begriff. Zudem ist das Zusammenspiel zwischen Lehrenden, Lernenden und dem Gegenstand wenig thematisiert, obwohl bspw. bekannt ist, dass sich die Umsetzung der einzelnen generischen Dimensionen zwischen den Lehrkräften unterscheidet (Kunter et al. 2005). Im Folgenden sollen deshalb diese drei Komponenten des Unterrichts – Lehrperson, Schülergruppe und Gegenstand –, welche den Aushandlungsprozess und damit die oben beschriebenen Merkmale in der jeweiligen Situation gestalten, näher beschrieben werden. Im Anschluss wird der Forschungsstand zu Schulartunterschieden im Unterricht und damit zusammenhängenden Merkmalen zusammengefasst.

3.2 Lehrende

Die Lehrperson nimmt eine besondere Rolle im Unterrichtsgeschehen ein. Durch die schulische Struktur wird sie beauftragt, die Schülerinnen und Schüler als „autoritatives Vollzugsorgan des Staates“ (Fend 2008, S. 106) anzuleiten und das Unterrichtsgeschehen zu planen, zu leiten und gelernte Inhalte zu überprüfen (Fend 2008). Zudem hat die Lehrperson die Aufgabe, die Unterrichtsinhalte auf ihren gegenwärtigen und zukünftigen Bedeutungsgehalt zu analysieren (Klafki 1996). Die Lehrperson stellt somit Lerngelegenheiten für die Schülerinnen und Schüler bereit. Hierbei stellt sich die Frage, was eine Lehrperson mitbringen sollte, um diesen Anforderungen gerecht zu werden.

Untersuchungen zur Frage, was eine gute Lehrperson ausmacht, haben eine vergleichsweise lange Tradition in der Bildungsforschung. In den Anfängen war besonders die Lehrerpersönlichkeit ein viel beforschtes Untersuchungsfeld (Rothland 2009). Dabei ging es um Personenmerkmale, die Unterschiede in der Art der erzieherischen Wirkung erklären sollten. Allerdings hatte diese Perspektive verschiedene Schwächen. Zum einen erwiesen sich die vermuteten Personeneigenschaften als entweder trivial oder als sehr komplex. Zum anderen blieb die Frage ihrer Auswirkungen und Ausformungen bei verschiedenen Unterrichtssituationen und Fachinhalten offen (Bromme 1997). Darüber hinaus war die Sicht auf den Beruf des Lehrenden stark vom Blickwinkel einer Beru-

fung, weniger von der Profession geprägt und bot somit wenig Potential das Handeln oder die Ausbildung der Lehrenden qualitativ zu prüfen oder zu optimieren (Rothland 2009).

Stärker handlungsorientiert ist die Untersuchung von Lehrerfertigkeiten (z. B. Klarheit der Lehrersprache oder Reaktionen auf Fehler) und deren Zusammenhang mit Schülerleistungen im Klassenzimmer aus einer behavioristisch geprägten Perspektive (Rothland 2009). Bei diesem als Prozess-Produkt-Paradigma bezeichnetem Ansatz ging man zunächst von einem direkten Zusammenhang zwischen Lehrerfertigkeiten und Schülerleistungen aus. Kritik an dieser isolierten Betrachtung führte teils zu einer programmatischen Erweiterung des Prozess-Produkt-Paradigmas (bspw. um Interaktionseffekte), teils zu seiner Ablehnung (Bromme 1997).

Die Erkenntnis, dass Lehrerhandeln stark kontextabhängig ist, rückte schließlich wieder die Lehrperson in den Blickpunkt der Analyse. Im Zuge der Professionalisierungsdebatten steht aber weniger die Persönlichkeit des Lehrenden im Fokus (Bastian et al. 2000), als vielmehr kognitive Merkmale, welche inzwischen als Grundlage des professionellen Lehrerhandelns betrachtet werden (Bromme et al. 2006). Erste Erkenntnisse über den Zusammenhang von Fachwissen und Unterrichtserfolg lieferten Fallstudien (bspw. Hashweh 1987; Roehler et al. 1987). Hier wurde deutlich, dass die curricularen Fachkenntnisse einer Lehrkraft dazu beitragen, Zusammenhänge und Sachverhalte im Curriculum hervorzuheben. Zudem beeinflussen fachliche Kenntnisse die Qualität von Erläuterungen und die Fähigkeit von Lehrenden, Schülerbeiträge zielführend in den Unterricht einzubinden (Bromme 1997). Auch der Grad der kognitiven Aktivierung der Lernenden scheint in Zusammenhang mit dem Fachwissen der Lehrperson zu stehen: In Unterrichtseinheiten, in denen eine Lehrperson weniger Fachwissen aufweist, stellt sie einfachere Fragen (low cognitive level), die Lernenden haben geringere Redeanteile und melden sich weniger (Carlsen 1987). In der empirischen Forschung liegt der Schwerpunkt im deutschsprachigen Raum mittlerweile in der Erfassung der Kompetenzen von Lehrenden sowie deren Zusammenhang mit Schülerleistungen. Hier hat sich das „Modell des Professionellen Handelns“ von Baumert und Kunter (2006) durchgesetzt, welches im folgenden Abschnitt näher beschrieben wird. Im Gegensatz zur Hypothese der persönlichen Eignung von Lehrenden, geht der Professionalisierungsansatz von der Erlernbarkeit des für den Lehrberuf relevanten Wissens und Könnens aus.

3.2.1 Professionelle Kompetenz von Lehrenden

In Anlehnung an professionelle Handlungsmodelle (Stichweh 1997) sowie aufbauend auf den Kernaussagen des National Board for Professional Teaching Standards (NBPTS 2002) führen Baumert und Kunter (2006) professionelles Handeln zurück auf spezifisches deklaratives und prozedurales Wissen (Professionswissen), professionelle Überzeugungen und Werthaltungen, motivationale Orientierungen sowie selbstregulative Fähigkeiten (Abbildung 5).

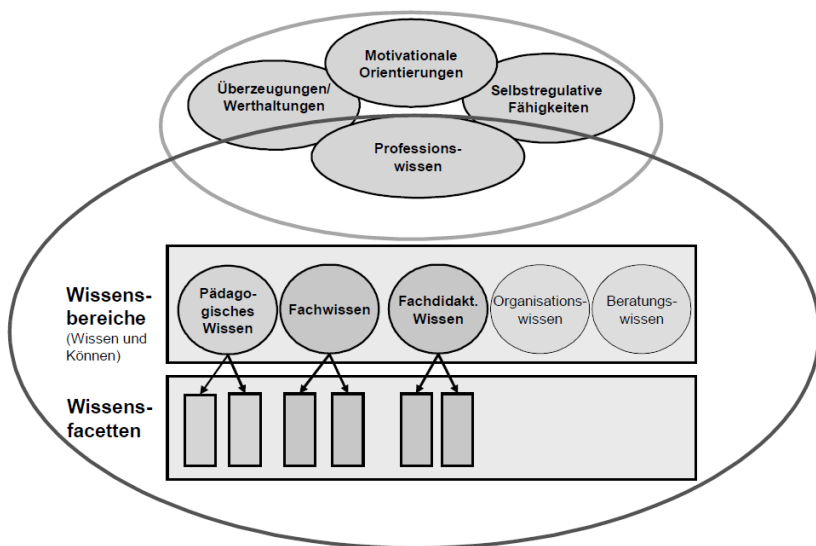


Abbildung 5. Das Kompetenzmodell von COACTIV.

Quelle: Baumert und Kunter (2006), S. 482.

Das *Professionswissen* wird in Anlehnung an Shulman (1986) differenziert in Fachwissen, fachdidaktisches Wissen und pädagogisch-psychologisches Wissen. Fachwissen hat sein Fundament in der akademischen Referenzdisziplin (bspw. Mathematik) und bezieht sich auf die Beherrschung des Schulstoffes, bspw. das Verständnis von mathematischen Prinzipien und Rechenoperationen. Um die Inhalte aufbereiten und vermitteln zu können, bedarf es darüber hinaus Wissens über das fachspezifische Lehren und Lernen (bspw. die didaktische Sequenzierung des Stoffes oder Darstellungsmöglichkeiten) sowie allgemeinen pädagogisch-psychologischen Wissens (bspw. Wissen über Unterrichtsmetho-

den oder diagnostische Grundlagen). Baumert und Kunter (2006) ergänzen zudem zwei Wissensbereiche, die über das professionelle Handeln in der Unterrichtssituation hinausgehen: das Organisationswissen sowie das Beratungswissen. Für alle Wissensbereiche wird angenommen, dass das „Wissen [...] zentrale Ressource ist, um Lerngelegenheiten in einem stabilen Ordnungsrahmen variationsreich, kognitiv herausfordernd und motivierend zugleich zu gestalten“ (Baumert et al. 2011, S. 14). Bezugnehmend auf Bromme (1997) und Schmidt und Boshuizen (1992) wird angenommen, dass deklaratives Fachwissen wie auch prozedurales Handlungswissen gleichsam bedeutsam und handlungsleitend sind, jedoch nicht immer auf direktem Wege. Vielmehr formt das professionelle Wissen die Grundlagen, „mit denen Unterrichtssituationen perceptiv strukturiert und damit auch interpretiert werden“ (Bromme 1997, S. 199). So können bspw. Kunter und Voss (2011) nachweisen, dass Teilbereiche der professionellen Handlungskompetenz von Mathematiklehrkräften spezifisch mit den Teildimensionen der Unterrichtsqualität zusammenhängen (z. B. fachdidaktisches Wissen mit dem Grad der kognitiven Aktivierung, allgemeinpädagogisches Wissen mit der Klassenführung).

Überzeugungen und Werthaltungen werden in einigen theoretischen Ansätzen als weitere Wissensdomäne definiert, meist als subjektives Wissen, das nicht von der Gesellschaft geteilt wird in Abgrenzung zu objektivem Faktenwissen (Turner et al. 2009). Im Modell von Baumert und Kunter (2006) sind Überzeugungen und Werthaltungen jedoch als eigene Kompetenzfacette definiert. Sie begründen diese Differenzierung in Anlehnung an Pajares (1992) und Op't Eynde et al. (2003): Im Unterschied zum Wissen müssen Überzeugungen und Werthaltungen weder frei von Widersprüchen sein, noch den Anforderungen der argumentativen Rechtfertigungen genügen. Es handelt sich um „überdauernde existentielle Annahmen über Phänomene der Objekte der Welt, die subjektiv für wahr gehalten werden, sowohl implizite als auch explizite Anteile besitzen und die Art der Begegnung mit der Welt beeinflussen“ (Voss et al. 2011, S. 235). Dies bezieht sich bspw. auf Überzeugungen über das Lernen im schulischen Kontext (subjektive Lerntheorien), subjektive Theorien über das Lehren des Gegenstandes oder Wertbindungen (value commitments). Grundlegend für den Zusammenhang zwischen Überzeugungen bzw. Werthaltungen und Handeln ist die These, dass das Handeln der Lehrperson im Unterricht vor allem von deren Vorstellungen darüber, wie sie unterrichten will, bestimmt wird (Wagner 2016). Im Zuge des zielgerichteten Handelns strukturieren Lehrende ihren Handlungsraum aktiv-kognitiv, d.h., die meist komplexen Situatio-

nen, denen sie sich gegenübersehen und die oft mehrdeutig, mehrdimensional, teilweise unvorhersehbar und immer kontextabhängig sind, werden fortlaufend analysiert, interpretiert und in bestimmter Weise rekonstruiert, um schließlich eine Handlungslinie zu entwickeln. Fives und Buehl (2012) konkretisieren diese Annahmen für das Zusammenspiel von Überzeugungen und Lehrerhandeln über drei Funktionen. Überzeugungen wirken 1. als Filter, welche die Wahrnehmung und die Interpretation von Informationen und Erfahrungen beeinflussen. Beispielsweise können sich Überzeugungen über angemessene Unterrichtsmethoden darauf auswirken, worauf Lehrkräfte bei der kognitiven Verarbeitung und Evaluierung von Informationen achten und wie sie diese Informationen interpretieren. Überzeugungen wirken 2. als Rahmen für Probleme und Aufgaben. Je nach Überzeugung über das Lehren und Lernen werden Lehrkräfte ein Problem unterschiedlich rahmen und einordnen und den Unterricht dementsprechend unterschiedlich planen und reflektieren. Zuletzt können Überzeugungen 3. die Planungen und Handlungen von Lehrkräften lenken (vgl. Dohrmann et al. 2019). Den Zusammenhang zwischen Überzeugungen und Unterrichtshandeln im Mathematikunterricht untersuchten unter anderem Voss et al. (2011) sowie Dubberke et al. (2008) in Mediationsanalysen. Lehrkräfte mit stark konstruktivistischen Überzeugungen realisieren im Unterricht ein höheres Potential zur kognitiven Aktivierung. Dieses Potential zur kognitiven Aktivierung ist wiederum vorteilhaft für die Leistungsentwicklung der Lernenden.

Die berufliche Tätigkeit von Lehrpersonen erfordert ein hohes Maß an Selbstständigkeit, Engagement und Leistungsfähigkeit. Deshalb zählen zur professionellen Kompetenz von Lehrpersonen neben fachbezogenem Wissen und Überzeugungen auch allgemeinere Kompetenzen der psychologischen Funktionsfähigkeit von handelnden Personen (Klusmann 2011). *Selbstregulative Fähigkeiten* sowie die *motivationale Orientierung* werden hier als zentrale Merkmale gesehen, um langfristig erfolgreich zu unterrichten (Baumert und Kunter 2011). Selbstregulation als Fähigkeit, „im beruflichen Kontext effektiv mit den eigenen Ressourcen haushalten zu können“ (Klusmann 2011, S. 277), scheint besonders angesichts hoher Frühpensionierungsraten und einer Häufung von Stresserkrankungen wichtig. Denn das Wohlbefinden wird als Erfolgskriterium insbesondere vor dem Hintergrund einer langen Berufslaufbahn erachtet (Kunter, Kleickmann et al. 2011; Vandenberghe und Hubermann 1999).

Das Kompetenzmodell von Baumert und Kunter (2006) beinhaltet zusammenfassend neben fachbezogenen kognitiven Kompetenzen (Wissen und Überzeugungen) auch überfachliche Aspekte (Motivation und Selbstregulation). Unter einer Professionsperspektive, wie sie das Modell professioneller Kompetenz von Baumert und Kunter (2006) einnimmt, wird davon ausgegangen, dass sich Lehrpersonen die Kompetenzen für ihre berufliche Tätigkeit im Verlauf der beruflichen Entwicklung in impliziten und expliziten Lernprozessen aneignen. Zunächst kommen die Lehramtsanwärter jedoch mit einem gewissen Grad an Wissen und Wertvorstellungen bzw. Überzeugungen in ihr Studium – durch die eigene schulische Ausbildung, aber auch durch die gesellschaftlich kulturelle Sozialisation (Gellert 1998; Geulen 2007; Hinsch 1979; Kunter und Voss 2011; Tanner 1993). Es ist naheliegend, dass bereits hier bestimmte Überzeugungen über das Lehren und Lernen an den verschiedenen Schularten vorliegen, wie sie in den Profilen der Schularten verankert sind (siehe Abschnitt 2.2). Die universitäre Ausbildung stellt dann das zentrale formale Lernangebot für die professionelle Kompetenz der Lehrkräfte dar. Es ist anzunehmen, dass gerade das Fachwissen sowie die Überzeugungen über das Lehren und Lernen in diesen formalen Lernsituationen konkretisiert werden (Kunter, Kleickmann et al. 2011; Terhart 2001). Durch die – auch weiterhin anhaltenden – Unterschiede in der Lehramtsausbildung in den Universitäten, aber auch im Referendariat sind bei diesen Dimensionen relevante Unterschiede zwischen den Lehrenden der verschiedenen Schularten zu erwarten, welche systematische Unterschiede im professionellen Handeln der Lehrkräfte erklären könnten. Das Fachwissen, das fachdidaktische Wissen sowie das pädagogische Wissen sind am Ende der Lehrerausbildung noch weitgehend kontextfreie und disziplinär strukturierte Kompetenzdimensionen (Jenßen et al. 2015). Erst mit der weiteren beruflichen Praxis werden diese stärker verknüpft, wobei sich hier Unterschiede zwischen den Lehrenden verschiedener Schularten abzeichnen (Blömeke 2019). Die Ausbildung der Lehrkräfte sowie potentielle Unterschiede in der Ausbildung zwischen den verschiedenen Lehramtsstudiengängen werden im Folgenden näher beleuchtet.

3.2.2 Erwerb professioneller Kompetenz

Die Ausbildung von Lehrkräften findet in Deutschland in zwei aufeinanderfolgenden Phasen statt: Einer überwiegend theoretischen Phase an Universitäten bzw. Pädagogischen Hochschulen und einer überwiegend praktischen Phase in

Ausbildungsschulen und Studienseminaren. Kompetenztheoretisch gesehen ist der Lehrberuf durch die Verknüpfung unterschiedlicher Wissensdimensionen gekennzeichnet (Fachwissen, Fachdidaktik und Erziehungswissenschaften; Blömeke 2009a), wie sie analog im Kompetenzmodell von Baumert und Kunter (2006) zu finden sind. Die drei Wissensdimensionen werden in der Regel separat in unterschiedlichen Fakultäten gelehrt. Hinter der Grundidee separierter Lerngelegenheiten stehen einerseits die Vorstellung unterschiedlicher Funktionen von Theorie und Praxis sowie andererseits die Annahme, dass eine qualitativ hochwertige Ausbildung in Fachwissenschaft, Fachdidaktik und Erziehungswissenschaften nur von den jeweiligen universitären Spezialisten geleistet werden kann (Blömeke 2009a). Erst in der zweiten, praktischen Phase der Ausbildung erfolgt eine handlungsorientierte Ausbildung. Praktische Routinen werden im Vorbereitungsdienst auf der Basis des zuvor erworbenen Wissens sorgfältig geplant und in Begleitung von erfahrenen Ausbildern lange eingeübt (Blömeke 2019).

Analog zum gegliederten Schulsystem in Deutschland führt auch die formale Struktur des Lehrberufes unterschiedliche Lehrämter, die jeweils verschiedene Stufen oder Zweige des Schulsystems abdecken. Damit gehen Unterschiede in der universitären Ausbildung sowie im praktischen Teil der Ausbildung einher, welcher in der Regel an einer der Schularten absolviert wird (Terhart 2008). Zum „Sachstand der Lehrerbildung“ legt die Kultusministerkonferenz mittlerweile jährliche Berichte vor (bspw. KMK 2017c) sowie Rahmenvereinbarungen für die Ausbildung und Prüfung der Lehrämter (KMK 1997) auf Basis derer Blömeke (2009b) Unterschiede in der Lehrerausbildung in den drei Wissensdomänen – Fachwissen, Fachdidaktik und pädagogisches Wissen – herausgearbeitet hat. Im Vergleich von angehenden Gymnasiallehrkräften mit Studierenden eines Grund-, Haupt- und Realschullehramts (GHR) sticht der Gesamtumfang des Studiums hervor, welcher für GHR-Studierende 7 bis 9 Fachsemester umfasst, für angehende Gymnasiallehrkräfte hingegen mindestens 9 Semester. Weiterhin unterscheiden sich die Semesterwochenstunden im Unterrichtsfach stark. GHR-Studierende belegen in ihren Unterrichtsfächern 30 bis 60 Semesterwochenstunden pro Fach, angehende Gymnasiallehrkräfte 60 bis 80 Semesterwochenstunden. Der stärkere Fokus auf das Fachstudium bei angehenden Gymnasiallehrkräften schlägt sich auch in der Ausprägung der Kompetenzen nieder. So zeigt sich im Rahmen der Studie „Mathematics Teaching in the 21st Century (MT21)“ (Blömeke et al. 2008) ein deutlicher Vorsprung im mathematischen und mathematikdidaktischen Wissen bei angehenden Gymna-

siallehrkräften gegenüber GHR-Studierenden. Angehende GHR-Lehrkräfte haben am Ende ihrer Ausbildung dagegen ein umfangreicheres erziehungswissenschaftliches Wissen erworben als angehende Gymnasiallehrkräfte (Blömeke et al. 2008; Blömeke 2009b). In den aktuellen Reformbestrebungen der Lehrerbildung wird die Ausbildungsdauer der verschiedenen Lehramtsstudiengänge zunehmend angepasst. Die Schwerpunktsetzung der Teildisziplinen scheint sich jedoch nicht zu vereinheitlichen. So folgert Terhart (2010, S. 15) für die Reformen in Nordrhein-Westfalen: „Unterhalb dieser Einheitlichkeit der Dauer werden in NRW hinsichtlich der Komposition, also des Verhältnisses der Anteile von Fachstudien, Fachdidaktik und Bildungswissenschaften etc. die Differenzen zwischen den Lehrämtern durch das neue Gesetz jedoch größer werden.“ Es ist also auch zukünftig mit Differenzen zwischen den Ausbildungsinhalten der verschiedenen Lehramtsstudiengänge zu rechnen.

Für die zweite Phase der Lehrerbildung mangelt es weiterhin an Analysen, welche die Unterschiede zwischen den Schularten herausarbeiten. Im Rahmen des großangelegten Projektes „Schulentwicklung und Lehrerarbeit“, welches Studienseminare in Hessen über mehrere Jahre evaluiert, kontrastiert Abs (2006) zumindest die Kompetenzentwicklung von Lehramtsanwärtern eines Grund-, Haupt- und Realschullehramts (GHR) und Gymnasialanwärtern in der zweiten Ausbildungsphase in Hessen. Abs (2006) findet deutliche Unterschiede zwischen den selbsteingeschätzten Kompetenzen zur Diagnostik und zur Unterrichtsvorbereitung dieser beiden Gruppen. Es zeigt sich, dass GHR-Anwärter ihre diagnostischen Kompetenzen höher einschätzen als es Gymnasialanwärter tun. Diese Unterschiede lassen sich in Regressionsanalysen zum Teil auf verschiedene Lerngelegenheiten während der zweiten Ausbildungsphase zurückführen: Für GHR-Anwärter scheinen besonders reflektierende Gespräche zu Unterrichtsstunden mit dem Mentor und Übungen im Seminar maßgeblich zur Kompetenzentwicklung beitragen. Für Gymnasialanwärter sind dagegen die schulischen Lernerfahrungen ausschlaggebend. Ähnliche Zusammenhänge zeigen sich bei der Kompetenz zur Unterrichtsvorbereitung: Für GHR-Anwärter zeigen sich die höchsten Zusammenhänge der Kompetenz zur Unterrichtsvorbereitung mit den Übungen im Seminar und der Reflexion mit dem Mentor. Für Gymnasialanwärter kann wiederum die praktische Lerngelegenheit in der Schule die größte Varianz in den Kompetenzunterschieden aufklären. Hieraus lässt sich folgern, dass GHR-Anwärter von mehr Lerngelegenheiten während der zweiten Phase der Ausbildung profitieren. Die Unterschiede in der eingeschätzten diagnostischen Kompetenz zwischen Gymnasialanwärtern und

GHR-Anwärtern verweist zudem auf die größere Schwerpunktlegung von Diagnostik und Förderung in der Ausbildung von GHR-Anwärtern. In der Ausbildung von Gymnasialanwärtern wird auf Kenntnisse der Diagnostik und Förderung evtl. kein Schwerpunkt gelegt, da darauf vertraut wird, dass durch das Elternhaus stark auf die Lernenden eingewirkt wird.

Neben den verschiedenen Schwerpunkten in den Lerngelegenheiten und Inhalten in der zweiten Phase der Lehrerbildung können auch die an den Ausbildungsschulen vorherrschenden Lehrkulturen die Überzeugungen und Wertvorstellungen der angehenden Lehrkräfte in je spezifischer Weise beeinflussen (Helsper und Kramer 2001). Hinzu kommen systematische Unterschiede in den Bildungszielen und -inhalten sowie in den Leitbildern der Schularten, welche in Abschnitt 2.2 herausgearbeitet wurden. Es ist anzunehmen, dass sich hierdurch analog zu den schulischen Lern- und Entwicklungsmilieus auf Schülerseite (Baumert et al. 2006b) auch für die Lehramtsanwärter schulartspezifische Sozialisations- und Professionalisierungsbedingungen herausbilden. Längsschnittuntersuchungen zu Einstellungen und Handeln von Lehrenden während des Berufseinstiegs scheint es keine zu geben. Für die fachlichen Kompetenzen von Berufseinsteigern in den Lehrerberuf finden sich hingegen über die Dauer von vier Jahren Unterschiede in der Kompetenzentwicklung bzw. der Vernetzung der Wissensdimensionen. So weisen Lehrkräfte an Gymnasien eine stärkere Vernetzung der drei Wissensdomänen auf als Lehrkräfte an GHR (Blömeke et al. 2014).

Zusammenfassend ist aufgrund der Unterschiede in der Gewichtung der fachlichen, fachdidaktischen und didaktischen Ausbildung der Lehrkräfte sowie in der Studiendauer davon auszugehen, dass sich die Kompetenzen in den verschiedenen Domänen sowie auch die Wertvorstellungen und Überzeugungen von Lehrkräften systematisch zwischen den Schularten unterscheiden. Auch für die zweite Ausbildungsphase gibt es, wie oben dargelegt wurde, Hinweise auf schulartspezifische Entwicklungen der Lehrenden in ihrem professionellen Wissen und Können sowie in ihren Überzeugungen. Da das Wissen und Können sowie Überzeugungen nachweislich im Zusammenhang mit dem Handeln der Lehrpersonen stehen, ist folglich von systematischen Unterschieden im unterrichtlichen Handeln der Lehrpersonen der verschiedenen Schularten auszugehen.

3.3 Gegenstand

Das Zusammenspiel von Lehren und Lernen findet immer in der Auseinandersetzung mit einem bestimmten sachlichen Gegenstand statt (Steindorf 1981). In Anbetracht des Umfangs und der Vielfalt des Wissbaren stellt sich jedoch die Frage, mit welchen Gegenständen die Lernenden konfrontiert werden sollten (Horlacher und Vincenti 2016). „Tausend Jahre, so hat man geschätzt, würden für einen Menschen nicht genügen, um die von der gesamten Menschheit zusammengetragenen Kenntnisse nachvollziehen zu können“ (Steindorf 1981, S. 79). Es gilt daher, aus der Gesamtheit von Wissen einzelne Aspekte – „das Bildungsgut, verstanden als geistiger Gemeinbesitz in einer Kulturgemeinschaft“ (Steindorf 1981, S. 75), – auszuwählen. Verbindliche Lehrpläne und Curricula sind Ausdruck der Selbstvergewisserung einer Gesellschaft, sie komprimieren, was an die jüngere Generation an kulturellen Beständen, an Wissen, Fähigkeiten und Fertigkeiten weitergegeben werden soll. Die Inhalte sind damit in bestimmten Abständen immer wieder Auseinandersetzungen über die enthaltenen Fächer, Themen und den jeweiligen Umfang, den diese einnehmen sollten, ausgesetzt (Giesecke 2002).

Die Auswahl der zu lernenden Inhalte stellt eine Vorentscheidung bezüglich der Lernmöglichkeiten des einzelnen Schülers bzw. der Schülerin dar, wie sie z. B. im Fremdsprachenangebot der Schulen sichtbar wird. Das Resultat des didaktischen Auswahlprozesses schlägt sich in der Stundentafel und vor allem im inhaltlichen Relief der Lehrpläne und Curricula nieder (Kiper und Mischke 2004, S. 39). Welchen Einfluss die gesellschaftlichen Entwicklungen und Bedeutungszuschreibungen an die Relevanz einzelner Bildungsaspekte haben, verdeutlichen die abweichenden Bildungsvorgaben für die Schularten. Am offensichtlichsten zeigt sich dies am Fächerkanon. Vergleicht man bspw. den Fächerkanon an Realschulen mit jenem für Gymnasien am Beispiel des Bundeslandes Bayern, gibt es neben den gemeinsamen Kernfächern für alle Schularten (wie z. B. Mathematik, Deutsch, Englisch, Biologie) deutliche Unterschiede. Eine Differenzierung in stärker praktisch orientierte Fächer, wie Hauswirtschaft, technisches Zeichnen oder Werken, legen die Lehrpläne und Curricula für Realschulen fest. Für Gymnasien ist der Fächerkanon stärker auf sprachliche, philosophische und kulturelle Fächer ausgerichtet (vgl. Staatsinstitut für Schulqualität und Bildungsforschung München 2001, 2010). Historisch lässt sich dies gut nachvollziehen, Realschulen entstanden überwiegend in Zeiten der Industrialisierung. Lernende sollten direkt auf Berufe in der Industrie

vorbereitet werden. Gymnasien sind durch die inhaltlichen Diskussionen verschiedener Bildungsphilosophen und Bildungstheoretiker geprägt und sollen auf das Studium vorbereiten (siehe Abschnitt 2.1). Genauere Analysen der Unterrichtsgegenstände verlangen einen tieferen Blick in die Unterrichtsinhalte und -ziele der einzelnen Fächer, verbindlich festgehalten in den Lehrplänen und Curricula. Hierfür gilt es, zunächst die Begriffsbestimmungen von „Lehrplan“ und „Curriculum“ zu klären, bevor deren Inhalte vergleichend zwischen den Schularten betrachtet werden.

3.3.1 Lehrpläne und Curricula – eine Begriffsbestimmung

Die Begriffe Curriculum und Lehrplan werden teilweise äquivalent verwendet. Allerdings sind die Begriffe mit verschiedenen Strömungen und Bedeutungszuschreibungen verbunden, weshalb es deren genauerer Definition bedarf (Horlacher und Vincenti 2016). Der Begriff „Curriculum“ hat die längere Tradition, wurde er doch bereits im barocken Mittelalter benutzt und insbesondere von Johannes Amos Comenius (1592–1670) für den Bildungsbereich fruchtbar gemacht (Kron et al. 2014, S. 193). In der historischen Verwendung bezeichnet der Begriff Curriculum einerseits einen individuellen Bildungsverlauf (*Curriculum vitae*) und andererseits ein institutionell organisiertes Lehrgangs- oder Schulangebot (*Curriculum scholasticum* oder *academicum*), das es zu durchlaufen gilt (Künzli 2009). Mit der Einführung der allgemeinen Schulpflicht im Laufe des 18. Jahrhunderts kam im deutschsprachigen Raum jedoch der Begriff des Lehrplans auf. Für die einzelnen Schularten wurden schrittweise verbindliche Lehrpläne erstellt (Dolch 1982; Horlacher und Vincenti 2016). Sie enthalten die Inhalte für eine Klassenstufe in einem bestimmten Fach. Sie sind in der Regel von einem staatlich berufenen Gremium für eine bestimmte Schulart zusammengestellt und damit legitimiert (Kron et al. 2014).

Als in den 1970er Jahren verschiedenen Studien große Defizite im deutschen Bildungssystem offengelegten (siehe Abschnitt 2.1.2), wurden – neben den auf organisatorische Veränderungen zielenden Debatten – auch die Bildungsinhalte stark diskutiert. Die Infragestellung der Lerngegenstände zielte dabei auf alle Schularten und jedes Lehrfach (Dolch 1982). Der Begriff Curriculum wurde in der Folge in Abgrenzung zu den überarbeitungswürdigen Lehrplänen – in Anlehnung an die angelsächsischen Curriculumdebatten – als rationalistisches Gegenprogramm zu dem vorwiegend geisteswissenschaftlichen, bildungstheo-

retischen und lehrpraktischen Diskurs über Didaktik und Lehrplan eingebracht (Künzli 2009). Dennoch weicht die Definition des Curriculumbegriffs inhaltlich von der angelsächsischen Tradition ab. Die deutsche Verwendung des Begriffs Curriculum hat vielfach nicht die thematische Reichweite des angelsächsischen Konzeptes. In der Tradition der Curriculum Studies in England stehen bspw. auch Fragen der Organisation des Curriculums sowie der gesellschaftlichen Auseinandersetzungen mit Schule und Prozessen des Lehrens und Lernens im Zentrum (Pinar 2009, S. 159 f). In Deutschland verebte die mit vielen Erwartungen und Hoffnungen eröffnete Debatte zur Etablierung des Begriffs und der Konzeptidee „Curriculum“ im Zuge breit gefächerter Überlegungen und Vorschläge zur „Curriculumrevision als Kern einer Schulreform“ (Böttcher 2003, S. 216) jedoch fast ebenso schnell, wie sie entstanden war. Beispielhaft zeigt sich dies an dem mehr als 900 Seiten umfassenden Handbuch der Curriculumforschung, welches im Titel die vielversprechenden Worte „Erste Ausgabe“ beinhaltet (Hameyer et al. 1983). Eine Fortsetzung oder Überarbeitung hat es jedoch nie gegeben. Gründe für das Scheitern sind unter anderem die überzogenen Erwartungen an die Planbarkeit von Bildungsprozessen, fehlende Aufarbeitung sozialwissenschaftlicher Theorien und mangelnde Kommunikation mit der Praxis. Die Curriculumforschung war somit nicht geeignet, den Erkenntnisbedarf der Bildungsplanung zu decken und geht in der Unterrichtsforschung auf (Knab 1983). Bildungsinhalte wurden daher weiterhin in Form von Lehrplänen festgehalten (Kron et al. 2014).

Erst in den 1990er Jahren findet der Begriff Curriculum wieder Einzug in die Bildungsdiskussion im Zuge der Debatte von Wirksamkeit politischer Steuerung. Bestand in den Bildungsverwaltungen der Bundesländer lange die Erwartung, dass die Lernenden die beschriebenen Bildungsziele und Inhalte auch in ein entsprechendes Können umsetzen, wurde diese Vorstellung durch die schlechten Ergebnisse deutscher Jugendlicher in internationalen Vergleichsuntersuchungen nachhaltig in Frage gestellt (Kron et al. 2014). Der Begriff Curriculum wurde im Zuge dessen als Abgrenzung zum Lehrplan, der Inhalte vorgibt, verwendet. Das Curriculum soll hingegen stärker zu erreichende Bildungsziele verbindlich festlegen. In der Literatur wird hier häufig von einem Paradigmenwechsel gesprochen – von der Inputsteuerung zur Outputsteuerung. Die Hoffnung ist, Curricula als Instrumente zum systematischen Kernkompetenzaufbau zu nutzen, aber dabei in hohem Maße auf die Autonomie und Professionalität der Lehrkräfte zu setzen (Böttcher 2003). Lernziele objektiv prüfbar zu machen, ist eine damit verbundene Prämisse, welche in Deutschland zur

Einführung verbindlicher Bildungsstandards geführt hat. Curricula sollen eine konsistente und zielkonforme Ausrichtung und Abstimmung der verschiedenen Orientierungs- und Steuerungselemente von Bildungsprozessen leisten (Künzli 2009). Die Entwicklungen folgen damit gesamtgesellschaftlichen Erwartungen an Transparenz und Rechenschaftslegung und den wachsenden methodischen Möglichkeiten, individuelle Lernleistungen und schulische Wirkungen objektivierend und vergleichend zu erfassen (Künzli 2009, S. 142).

Die Definition der Begriffe „Curriculum“ und „Lehrplan“ verdeutlicht, dass verschiedene Strömungen Inhalt und Umfang beider Begriffe prägen. Während der Begriff Lehrplan mit den Erfahrungen und Vorstellungen der deutschen Schul-, Bildungs- und Wissenschaftskultur angereichert ist, ist die heutige Bedeutung des Begriffs Curriculum zunehmend durch angelsächsische und amerikanische Verwendungszusammenhänge beeinflusst (Künzli 2009, S. 134).

3.3.2 Lehrpläne und Curricula der Schularten

Inzwischen werden je nach Bundesland Lehrpläne oder Curricula oder teils beides ausgegeben, nachdem sich in einem Beschluss der KMK alle Bundesländer dazu verpflichtet haben, die nationalen Bildungsstandards zu übernehmen, zu implementieren und anzuwenden (KMK 2003). Einen Überblick der geltenden Lehrpläne und Curricula bietet die Lehrplandatenbank der KMK (KMK 2018). Das Bundesland Hessen hat bspw. ab dem Jahr 2011 die Vorgaben stückweise überarbeitet und Kerncurricula für die einzelnen Fächer und Bildungsgänge ausgegeben, welche sich an den Bildungsstandards orientieren (Hessisches Kultusministerium 2011c). Die Inhalte sollen von den Schulen selbst als schuleigene Curricula gefüllt werden (Hessisches Kultusministerium 2011a, S. 6):

„Das neue Kerncurriculum für Hessen bildet die Grundlage für die Entwicklung schulinterner Curricula. Die Kompetenzformulierungen und Inhaltsfelder sind auf die wesentlichen Aspekte, den Kern eines Faches, fokussiert und stellen die zentralen Ankerpunkte für die Ausgestaltung der jeweiligen Fachcurricula dar. Dies bietet zum einen eine Orientierung für schulische Entwicklungs- und Planungsprozesse, eröffnet aber auch Gestaltungsräume, um Interessen und Neigungen der Lernenden situationsbezogen in die curriculare Planung einzubeziehen bzw. regionale und schulspe-

zifische Besonderheiten zu berücksichtigen. Damit wird das schuleigene Curriculum zum Bindeglied zwischen dem Kerncurriculum und der individuellen Unterrichtsgestaltung.“

Hierfür verweist das Hessische Kultusministerium wiederum auf die letzte Version der Lehrpläne (Hessisches Kultusministerium 2011c). Inhaltlich unterscheiden sich die Curricula zwischen den einzelnen Schularten kaum. Ein Dokumentenabgleich zwischen den Kerncurricula Mathematik für die Sekundarstufe I offenbart ca. 70 abweichende Worte zwischen den Vorgaben für das Gymnasium und für die Realschule (von insgesamt ca. 7130 Wörtern pro Dokument). Zudem sind die Aussagen auf sehr abstraktem Niveau (Hessisches Kultusministerium 2011a, 2011b). Als Bildungsziel wird für Gymnasiasten bspw. angegeben (Hessisches Kultusministerium 2011a, S. 17): „Die Lernenden [...] analysieren, erläutern und begründen mathematische Aussagen und Verfahren.“ Lernende an Realschulen haben das gleiche Bildungsziel, mit der Ausnahme, dass sie mathematische Aussagen nur analysieren, jedoch nicht erläutern und begründen können sollen (Hessisches Kultusministerium 2011b).

Die Orientierungshilfe für die Einzelschule ist dementsprechend gering. Es stellt sich daher die Frage, ob die Schulen für ihr schuleigenes Curriculum verstärkt auf die alten Lehrpläne zurückgreifen, oder tatsächlich schuleigene Spezifika und Inhalte konkretisieren. Ersteres würde systematische Unterschiede zwischen den Schularten fortführen, letzteres würde hingegen die Varianz auf Schulebene vergrößern. Dies zu untersuchen wäre angesichts der Schulstrukturdebatte sehr aufschlussreich. Analysen im Rahmen der PISA-2000-Studie lassen zumindest für die früheren Lehrpläne auf systematische Unterschiede in den Inhalten der Fächer Mathematik und Naturwissenschaften zwischen den Schularten schließen. Untersucht wurde die curriculare Validität der PISA-Aufgaben unter anderem, indem Experten befragt wurden ob bzw. in welcher Klassenstufe PISA-Aufgaben der verschiedenen Kompetenzniveaus (siehe hierzu die Beschreibung der Kompetenzstufen Klieme, Neubrand et al. 2001, 147 ff) von den Lernenden der Haupt-, Real- oder gymnasialen Schulen lösbar wären (Artelt et al. 2003). Im Bereich Mathematik zeigt sich, dass die Kompetenzen zum Lösen der Aufgaben in Hauptschulen meist später erworben werden als in Realschulen. Ähnlich verhält es sich im Vergleich von Realschulen und Gymnasien. Noch deutlicher zeigen sich Unterschiede im Mathematiklehrplan zwischen den Schularten in den Analysen zu PISA 2003. Hier wurde auf die Unterscheidung der Kompetenzniveaus verzichtet. Stattdessen wurde

der Anteil an Aufgaben berichtet, welche in den verschiedenen Schularten bis zur neunten Klasse laut den Vorgaben in Lehrplänen behandelt wurden. In Hauptschulen werden nach Angaben der Experten 87,6 % der Aufgaben bis zur 9. Klasse behandelt. In integrierten Gesamtschulen sind es 91,4 %. Realschulen und Gymnasien haben den Stoff zu 90,2 % bzw. 93,1 % behandelt (Blum et al. 2004, S. 64).

3.3.3 Schulische Lerngelegenheiten

Die Ergebnisse des vorangegangenen Abschnitts werfen nun die Frage auf, was den Lernenden tatsächlich im Unterricht gelehrt wurde und welche Bedeutung Bildungsinhalte und -umfang für die Leistung der Lernenden haben. In der empirisch analytischen Unterrichtsforschung wird diese Frage traditionell unter dem Begriff der Lerngelegenheit diskutiert. Durch den starken Einfluss der meist englischsprachigen Schuleffektivitätsforschung wird häufig auch der Begriff „Opportunity to learn“ (OTL) verwendet (Klieme 2018). Der Begriff OTL wurde in den 1960er Jahren von Carroll (1963) eingeführt. Er betont den Aspekt der Lernzeit: „the amount of time allowed for learning, for example by a school schedule or program“ (Carroll 1989, S. 26). In der heutigen Verwendung ist der Begriff breiter und umfasst Aspekte der Qualität wie auch der Quantität von Lerngelegenheiten. Neben der Lernzeit werden auch Lerninhalte als Quantitätsaspekt mit den Schülerleistungen in Bezug gesetzt (Kurz 2011). Qualität und Quantität des Unterrichts werden in dieser Arbeit getrennt betrachtet. Die Quantität des Unterrichts ist durch die Spezifikation der Inhalte und des Unterrichtsumfangs hier als Gegenstand des Unterrichts spezifiziert. Die Unterrichtsqualität ist hingegen Teil der Unterrichtsprozesse, die in Abschnitt 3.1.3 diskutiert werden.

Inzwischen ist OTL zu einem breiten Forschungsfeld geworden. Vor allem in international vergleichenden Studien wie TIMSS und PISA werden Lehrende sowie Lernende regelmäßig befragt, welche Konzepte im Unterricht behandelt wurden und wie viele Unterrichtsstunden sie in bestimmten Fächern haben. In den meisten Studien zeigen sich Lerninhalte und Lernzeit als relevante Faktoren zu Erklärung von Schülerleistungen (eine Zusammenfassung des Forschungsstandes geben Kuger 2016; Scheerens et al. 1997; Scheerens 2017). Bewährt hat sich bei der Erklärung von Schülerleistungen eine Mehrebenenperspektive auf die Bildungsmöglichkeiten der Lernenden, wissend, dass Bil-

dungssysteme, Schulen, Lehrende und die Lernenden selbst die Lernmöglichkeiten und Erfahrungen des Einzelnen beeinflussen (Ditton 2000). Bezogen auf die Lernzeit und die Lerninhalte haben Schmidt und Cogan (1996) im Rahmen der TIMSS-Studie ein dreistufiges Modell vorgeschlagen, welches inzwischen international wie auch in deutschsprachigen Publikationen häufig Verwendung findet. Im Zentrum dieses Modells steht die Transformation eines von der Bildungsadministration *intendierten Curriculums*¹⁰, das nationale und regionale Bildungsziele vorgibt. Auf Ebene der Schule und Klasse wird das intendierte Curriculum ergänzt und konkretisiert, durch die Ziele der Schulen und Lehrpersonen. Daraus resultiert das *implementierte Curriculum*. Das *erreichte Curriculum* meint schließlich die Leistungsresultate der Lernenden auf individueller Ebene (Ditton 2000; Schmidt und Cogan 1996). Das intendierte Curriculum gibt damit den Handlungsspielraum in der inhaltlichen und zeitlichen Gestaltung des Unterrichts vor und erhöht damit die Wahrscheinlichkeit, dass bestimmte Inhalte gelehrt werden und andere Inhalte ausgeschlossen werden. Zudem wird der Umfang der Unterrichtsfächer bestimmt. Das implementierte Curriculum ist inhaltlich stärker auf die Ziele der Schule und die Lernenden zugeschnitten, umgesetzt durch den Lehrenden mit seinen je eigenen Fähigkeiten und Vorstellungen von Unterricht (Ditton 2000). Auch die vorgegebene Lernzeit wird auf Schul- und Klassenebene transformiert. Krankheit des Lehrpersonals, Schulausflüge und Projekttag beeinflussen die tatsächlich unterrichtete Zeit eines Faches oder Lerninhaltes. Darüber hinaus wirken konkrete Prozesse auf Klassenebene, wie die Klassenführung oder die Disziplin, auf die Lernzeit. Das erreichte Curriculum ist schließlich von individuellen Faktoren der Lernenden mitbestimmt. Motivation, individuelle Bildungsziele und das Leistungsvermögen spielen hierbei eine entscheidende Rolle (Kuger 2016).

Wegen der Transformation von Lehrplan und Curriculum auf Ebene der Schule und des Unterrichts ist es von besonderem Interesse, welche Inhalte, in welchem Umfang und in welcher Tiefe tatsächlich angeboten wurden. Der Beantwortung dieser Frage widmen sich in Deutschland überwiegend Leistungsvergleichsstudien. In PISA 2012 wurde bspw. abgefragt, welche Aufgabenformen und -inhalte im Mathematikunterricht durchgenommen wurden und in welchem Umfang (Schiepe-Tiska et al. 2013). Für die 15-jährigen in Deutschland fällt

10 Der Curriculumbegriff ist aus dem englischsprachigen Kontext entlehnt (siehe Abschnitt 3.3.1). Rekurrierend auf obige Begriffsbestimmungen von Lehrplan und Curriculum wäre jedoch der Begriff des Lehrplans durch den stärkeren Inhaltsbezug zu bevorzugen.

hierbei auf, dass Aufgaben ohne Anwendungsbezug im Mathematikunterricht eine deutlich größere Rolle spielen als anwendungsbezogene Aufgaben: „67,7 % der Jugendlichen geben an, dass Aufgaben, bei denen eher das technische Anwenden von Algorithmen wie zum Beispiel das Lösen einer Gleichung im Vordergrund steht, ‚häufig‘ im Unterricht vorkommen. Dagegen wählen nur 22,6 % diese Einordnung für die Aufgabe zum Berechnen eines Kaufpreises mit Mehrwertsteuer“ (Schiepe-Tiska et al. 2013, S. 137). Dabei wird gerade dem Anwendungsbezug – besonders von Seiten der Lehrkräfte – eine bedeutende Rolle in der Kompetenzentwicklung zugeschrieben (Baumert et al. 2004, S. 324). Inzwischen hat sich jedoch gezeigt, dass es verschiedene Facetten des anwendungsbezogenen Unterrichts gibt, welche im leistungsfördernden Potential variieren. Für den naturwissenschaftlichen Unterricht haben dies bspw. Schiepe-Tiska et al. (2013) herausgearbeitet. Sie unterscheiden vier Formen des forschend-entdeckenden Unterrichts, wobei sie dem epistemischen Unterricht, welcher in Deutschland im internationalen Vergleich selten praktiziert wird, die größte Bedeutung für die Leistungsentwicklung zuschreiben. Für den Mathematikunterricht regten insbesondere die Ergebnisse der TIMS-Videostudie Ende der 1990er Jahr in Deutschland eine intensive Auseinandersetzung mit dem kognitiven Anspruchsniveau des Mathematikunterrichts an. Klieme, Schümer et al. (2001) verglichen im Rahmen dieser Studie den Unterricht deutscher und japanischer Mathematikklassen und kamen zu dem Ergebnis, dass japanische Unterrichtsstunden stärker von individuellen, aktiven und problemorientierten Arbeitsphasen geprägt sind. Zudem sind die Aufgaben offener und komplexer als im deutschen Unterricht. Japanische Lernende zeigen im internationalen Vergleich kontinuierliche Spitzenleistungen in Mathematik. Ausgehend von diesen Befunden untersuchen Bohl et al. (2013) in einer qualitativen Studie die Aufgabenkultur im Mathematikunterricht in verschiedenen Schulklassen anhand von videografierten Unterrichtsaufzeichnungen und Befragungen von Mathematiklehrkräften. Sie arbeiteten bedeutsame Merkmale im Einsatz und Umgang mit mathematischem Aufgabenmaterial heraus, bspw. in der Einführung, der Lernbegleitung und der Besprechung von Aufgaben, sowie in den Fragearten (Reproduktionsfragen, Faktenfragen, Begründungsfragen). Einen direkten Bezug zum Lernen oder der kognitiven und motivationalen Entwicklung der Lernenden stellen sie nicht her. Sie können aber deutliche Unterschiede zwischen den Klassen der einzelnen Schularten (Hauptschule, Realschule und Gymnasium) finden, die vermuten lassen, dass stärker kognitiv herausfordernde Aufgaben mit einer höheren Mathematikleistung einhergehen.

Die Ergebnisse zu den Schulartunterschieden in der Aufgabenkultur werden ausführlicher in Abschnitt 3.5 beschrieben.

Stärker beforscht als die Unterrichtsinhalte und das Anspruchsniveau der Aufgaben ist die Unterrichts- und Lernzeit, welche die Lernenden für die einzelnen Fächer oder Themen nutzen können. Eine Reihe von Forschungsarbeiten weist nach, dass sich „time on task“, also die aufgabenbezogene Nutzung der Lernzeit, positiv auf den Lernerfolg auswirkt (zusammenfassend siehe Fredrick und Walberg 2014; Hattie 2009). Die Lernzeit bezieht sich zum einen auf die Beschulungsjahre in einem bestimmten Fach und auf die intendierten, durch Lehrpläne festgeschriebenen Unterrichtsstunden. In den meisten Studien zeigt sich ein positiver Zusammenhang der Beschulungsjahre bzw. Stunden mit der Leistung der Lernenden. Diese Erkenntnis ist evident und hat im deutschen Bildungssystem bereits zu Anpassungen im Einschulungsalter geführt, indem das Einschulungsalter herabgesetzt wurde, damit die 15-Jährigen aufgrund dieses gewonnenen Beschulungsjahres besser in PISA abschnitten (Bellenberg und Klemm 2011). Über die gesamte Schulzeit hinweg sollte diese Änderung allerdings keinen Effekt haben, da die Gesamtschulzeit nicht verlängert, in Gymnasien teilweise sogar gekürzt wurde (G8). Tiefergreifende Analysen zur Lernzeit haben zudem ergeben, dass es nicht nur auf die Brutto-Lernzeit ankommt, sondern mehr auf die optimale Nutzung der Unterrichtsstunden. Ausfall von Unterricht und individuelle Fehlzeiten erweisen sich als Parameter, die sich in hohem Maße negativ auf die schulische Leistung auswirken: „In leistungsstarken Klassen wird die komplette Unterrichtszeit für die Arbeit am Unterrichtsthema ausgenutzt. Dies bezieht sich nicht nur auf einen zügigen Unterrichtsbeginn und einen Verzicht auf ein vorzeitiges Ende, sondern auch auf das Vermeiden von Unterrichtsstörungen“ (Blömeke und Müller 2009, S. 240).

Die schulischen Lerngelegenheiten, bezogen auf die dargebotenen Inhalte und den zeitlichen Umfang, stellen zusammenfassend eine bedeutende Komponente im schulischen Lernen der Schülerinnen und Schüler da. Dies bezieht sich sowohl auf die Ebene des intendierten Curriculums als auch auf die Ebene des realisierten Curriculums. Unterschiede in Inhalten und Umfang auf Ebene des intendierten Curriculums wurden in diesem Abschnitt bereits vergleichend zwischen den Schularten thematisiert. Erkenntnisse aus der Vergleichenden Forschung zu Aufgabenkulturen in den Schularten oder Faktoren, welche die optimale Nutzung der Unterrichtszeit auf Klassenebene beeinflussen, werden in Abschnitt 3.5 beschrieben.

3.4 Schülergruppe

Die Eigenschaften, welche die Lernenden in eine Klasse oder Schule mitbringen, führen zusammengenommen zu einer ganz bestimmten Zusammensetzung der Schülergruppe. Diese Zusammensetzung formiert sich dabei innerhalb eines hierarchischen Systems: Einzugsgebiet, Schule, Klasse und Lerngruppe (Dreeben und Barr 1988). Im deutschen Bildungssystem zeigt sich zudem, dass die auf Schülerleistungen basierende Allokation der Lernenden an verschiedene Schularten sowie soziale Segregationsprozesse zu einer schulartspezifischen Zusammensetzung der Schülerschaft führen (Klinge 2016; Maaz 2017). Auch hinsichtlich affektiver Merkmale finden verschiedene Studien Unterschiede zwischen den Schularten, bspw. bezogen auf Lesefreude, das Selbstkonzept oder die Motivation (bspw. Kunter 2005; Pekrun und Zirngibl 2004; Stäbler 2017). Somit ergeben sich zwischen den Klassen der verschiedenen Schularten unterschiedliche Ausgangsbedingungen für die Gestaltung des Unterrichts (die Unterschiede zwischen den Schularten bezogen auf die Schülergruppe sind ausführlicher beschrieben in Abschnitt 3.5). In Deutschland wie auch international haben verschiedene Forschergruppen die Zusammenhänge zwischen der Zusammensetzung der Schülergruppe (Komposition) und den Entwicklungsmöglichkeiten des Einzelnen untersucht. Besonders für die Leistungskomposition konnten Effekte der Lerngruppe auf die individuelle Leistung bzw. Leistungsentwicklung gefunden werden (Baumert et al. 2006b; Dumont et al. 2013; Scharenberg 2012). Auch für die soziale Zusammensetzung der Gruppen finden sich in einigen Studien Effekte der Lerngruppe auf die individuelle Entwicklung. Diese fallen jedoch deutlich kleiner aus. Zudem sind die Indikatoren der leistungsbezogenen Komposition sowie der sozialen Komposition in hohem Maße konfundiert. Angenommen wird, dass die verschiedenen Kompositionsmerkmale zum einen direkt mit der individuellen Leistungsentwicklung zusammenhängen, aber auch Effekte auf Schülermerkmale wie Bildungsaspirationen, Leistungsmotivation und Selbsteinschätzungen vorliegen, die wiederum mit der Leistungsentwicklung verbunden sind (Baumert et al. 2006b; Wilkinson et al. 2000).

Die Untersuchung von Kompositionseffekten trägt wesentlich dazu bei, die Bedeutung der Lernumgebung für das schulische Lernen der Schülerinnen und Schüler zu ergründen. Von besonderem Interesse, besonders für die Unterrichtsforschung, ist es, nicht nur die Existenz von Kompositionseffekten nachzuweisen, sondern die dahinterliegenden Mechanismen zu identifizieren (Wil-

kinson 2002). Anzunehmen ist, dass die Effekte vor allem durch Prozesse auf proximaler Klassenebene vermittelt werden (Bellin 2009). „Vermittelt“ bedeutet hier, dass Effekte der Klassenzusammensetzung auf Lernergebnisse von Schülerinnen und Schülern durch andere Klassenmerkmale indirekt geleitet werden und durch sie (teilweise) erklärbar sind. Gerade aus diesen Wirkmechanismen könnten sich relevante Unterrichtmerkmale in Zusammenhang mit der Gruppe der Lernenden ableiten lassen.

Die in der Forschungsliteratur diskutierten Wirkmechanismen lassen sich im Wesentlichen vier Erklärungsansätzen zuordnen. Der erste Ansatz argumentiert über die Anpassung von Normen, Werten und Einstellungen. Hier wird angenommen, dass schulische Normen, Werte und Einstellungen das Lernen der Schülerinnen und Schüler lenken und sich die Lernenden gegenseitig über Interaktionen und Gruppendynamiken in ihren Einstellungen und Fähigkeiten beeinflussen (Dreeben und Barr 1988). Diese Annahmen sind in der Literatur bislang sehr prominent, dennoch wenig theoretisch fundiert. Deshalb wird in dieser Arbeit dieser Erklärungsansatz ausführlicher beschrieben und theoretisch hergeleitet (Abschnitt 3.4.2). Ein vorangestellter Exkurs reflektiert speziell die häufige Erklärung von Effekten der sozialen Komposition über die Normen, Werte und Einstellungen der Lernenden (Abschnitt 3.4.1). Ein zweiter Ansatz zur Erklärung von Kompositionseffekten argumentiert nicht über die interaktiven Prozesse, sondern nimmt die gebündelten Fähigkeiten, Einstellungen und Verhaltensweisen als Ressource und Grundlage einer je spezifischen Lernumgebung (Abschnitt 3.4.3). Der dritte Ansatz argumentiert über eine komparative Funktion der Gruppe für die Entwicklung des Fähigkeitskonzepts (Abschnitt 3.4.4). Der vierte Ansatz fokussiert auf die Lehrenden. Hier wird angenommen, dass Lehrende auf die Zusammensetzung der Lerngruppe bewusst reagieren oder indirekt durch bestimmte Erwartungen gegenüber dem Verhalten oder den Leistungen der Lerngruppe agieren (Abschnitt 3.4.5). Im Folgenden werden diese Prozesse näher beschrieben. Eine Trennung nach Wirkmechanismen verschiedener Kompositionsmerkmale (wie Leistungskomposition oder sozioökonomische Komposition) kann dabei nicht immer aufrechterhalten werden, da die Mechanismen nicht isoliert voneinander verlaufen, sondern gleichzeitig und wechselwirkend (Nash 2003). Beispielsweise können Einflüsse der Komposition bezüglich Normen, Werten und Einstellungen die Leistungsbereitschaft und somit die tatsächliche Leistung beeinflussen, welche ihrerseits als Kompositionsmerkmal die Lernprozesse im Unterricht beeinflussen kann.

3.4.1 Exkurs: Soziale Herkunft und schulische Normen, Werte und Einstellungen?

Werden Kompositionseffekte der sozialen Zusammensetzung diskutiert, werden diese häufig über die Normen, Werte und Einstellungen der Lernenden als einer der möglichen Wirkmechanismen erklärt. Hierbei liegt implizit die Annahme zugrunde, dass Lernende mit hohem sozioökonomischem Status über lernförderlichere Normen, Werte und Einstellungen verfügten als sozioökonomisch schwächer gestellte Lernende (bspw. Rjosk et al. 2014). Lernende aus privilegierten Familien verfügten demnach über positivere Einstellungen zur Schule, maßten der Bildung eine höhere Bedeutung bei und zeigten eine höhere Leistungsbereitschaft. Defizite in der Aufmerksamkeit, der Disziplin und dem Interesse werden wiederum eher Lernenden aus bildungsfernen Familien zugeschrieben. Es bleibt jedoch weitgehend ungeklärt, warum Lernende mit hoher sozioökonomischer Stellung lernförderliche Einstellungen haben sollten, Lernende aus sozioökonomisch schlechter gestellten Familien hingegen negative Einstellungen.

Diese als defizitär anmutende Hypothese bezüglich bildungshinderlicher Normen, Werte und Einstellungen von Personen mit niedrigem sozioökonomischem Hintergrund soll in diesem Abschnitt zunächst hinterfragt werden. Mit Bezug zu Bourdieu können Argumente für diesen postulierten Zusammenhang gefunden werden. Es werden aber auch Gegenargumente aufgeführt mit der Schlussfolgerung, Normen, Werte und Einstellungen von Lernenden bevorzugt direkt als Kompositionsmerkmal zu untersuchen.

Nach Bourdieu haben Einzelne und soziale Gruppen unterschiedliche Möglichkeiten des Handelns aufgrund von akkumuliertem, vererbtem und übertragbarem Kapital. Unterschieden werden hierbei das kulturelle, das ökonomische sowie das soziale Kapital. Für die bildungsbezogenen Normen, Werte und Einstellungen ist das kulturelle Kapital hervorzuheben. Das kulturelle Kapital umfasst neben objektiviertem Kapital (bspw. Bücher, Musikinstrumente oder Kunstobjekte) und institutionalisiertem Kapital (bspw. Bildungsabschlüsse der Eltern) – auch inkorporiertes Kapital in Form von dauerhaften Dispositionen (Bourdieu 1983). Bourdieu geht davon aus, dass jede Familie ihren Kindern auf überwiegend indirekten Wegen ein System impliziter und tief verinnerlichter Werte vermittelt, welches die Einstellungen zur Schule und zum Wert von Kultur entscheidend beeinflusst. Dieses Kapital muss vom Individuum selbst

verinnerlicht werden und kann nicht kurzfristig weitergegeben werden, bspw. kann es nicht wie ökonomisches Kapital verschenkt werden. Stattdessen kann es implizit über den Prozess der sozialen Vererbung weitergegeben werden. Das greifbarste Beispiel hierfür ist die Sprache einer bestimmten Klasse (Bourdieu 1983). Bourdieu geht damit von einer Art kulturellem Erbe aus, das nach sozialen Klassen bzw. Milieus variiert und für die Varianz der schulbezogenen Einstellungen und Handlungen von Kindern verantwortlich ist (Bourdieu 2001).

Als negativ bewertet werden die variierenden Einstellungen und Handlungen der Lernenden mit niedrigerem sozioökonomischen Hintergrund dann, wenn innerhalb der institutionalisierten Bildung die Normen, Werte und Einstellungen der dominierenden Klassen bzw. Milieus als die Gewünschten erachtet werden und somit die Normen, Werte und Einstellungen der anderen Klassen bzw. Milieus von dem Gewünschten abweichen (Bourdieu 2001). Dies ist der Fall in der institutionalisierten Bildung, welche geprägt ist von den Normen, Werten und Einstellungen der sozial und ökonomisch höher gestellten Bevölkerung, beispielsweise werden in Deutschland Lehrpläne und somit Bildungsinhalte von Akademikern festgelegt. Ebenso wird der Unterricht von Akademikern gehalten, welche nicht zuletzt eine sprachliche Norm setzen, aber auch ihre inkorporierten kulturelle Werte und Einstellungen als allgemeingültig annehmen: „Einstellungen und Fähigkeiten, die den Angehörigen der kulturellen Klasse nur deshalb als natürlich und selbstverständlich einforderbar erscheinen, weil sie die Kultur [...] dieser Klasse ausmachen.“ (Bourdieu 2001, S. 41).

An diese Auffassung knüpft auch die Schulkulturtheorie der Forschergruppe um Werner Helsper an. Es wird angenommen, dass die schulischen Akteure – mit dem Fokus auf Schulleitung und Lehrpersonal – einen pädagogischen Sinnentwurf konstruieren. Dies geschieht unter Rückgriff auf übergreifende soziale und institutionelle Strukturen und die je eigenen milieuspezifischen Kontexte der Akteure (Helsper et al. 2005; Helsper und Kramer 2001). Damit einher geht auch ein pädagogischer Habitus- und Zielentwurf, welchem die Lernenden näherungsweise entsprechen müssen, um durch die pädagogischen Akteure Anerkennung zu erfahren (Helsper et al. 2005). Passt der Habitus der Lernenden hingegen nicht zum schulischen Entwurf, ist mit schulischen Problemen zu rechnen, welche bis zum Abbruch der Laufbahn an der Schule führen können. Anschauliche Beispiele für Passungsprobleme zwischen dem primären, im

familiären Umfeld erworbenen Habitus und den schulischen Sinnstrukturen sowie den daraus resultierenden Anerkennungsdefiziten und Gegenreaktionen finden sich bei Böhme (2000). Sie arbeitet heraus, dass es oppositionellen Lernenden mit abweichendem Verhalten weniger um eine Abwertung des schulischen Sinnangebotes geht, sondern durch das abweichende Verhalten die Anerkennung als gleichberechtigter Interaktionspartner erreicht werden soll. Somit haben Lernende mit einem primären Habitus, der den pädagogischen Zielen der Schule entspricht, größere Erfolgschancen und können schnell als Vorzeigeschüler gelten (Helsper et al. 2005; Helsper und Böhme 2008).

In Rückbezug auf Bourdieu ist anzunehmen, dass Anerkennungs- und Anpassungsprobleme eher bei Lernenden aus niedrigeren sozialen Klassen zu erwarten sind. Trotz schulspezifischer Sinnentwürfe gibt es im Bildungssystem eine grundlegende Übereinstimmung in schulischen Erziehungs- und Bildungsentwürfen bezogen auf die Leistungs- und Erfolgsorientierung bzw. die Zielstrebigkeit (Helsper et al. 2005), resultierend aus dem Habitus der dominierenden Klassen (Bourdieu 2001). Lernende aus höheren sozialen Klassen bzw. Milieus stehen damit der Kultur der Schule eher nahe. Ihre Passung in Sprache und Neigungen wird als Begabung interpretiert, wohingegen die Lernenden aus niedrigeren sozialen Klassen bzw. Milieus sich mühsam an die dominierende Kultur anpassen müssen, was nicht zuletzt als fehlende Begabung ausgelegt wird.

Die Abwertung der Normen, Werte und Einstellungen der benachteiligten Schichten bzw. Milieus und die damit einhergehenden exkludierenden Mechanismen können sich nach Bourdieu noch in anderer Weise negativ auf die schulischen Einstellungen der betroffenen Lernenden auswirken. Basierend auf den Erfolgen bzw. Misserfolgen der Kinder und Jugendlichen aus der Nachbarschaft und dem sozialen Umfeld bilden die Eltern eine Art „intuitiver Statistik“ (Bourdieu und Passeron 1971, S. 22). Da Ansprüche und Forderungen in Form und Inhalt nach Bourdieu durch die objektiven Bedingungen bestimmt werden, haben Eltern aus benachteiligten sozialen Schichten geringere bildungsbezogene Ansprüche und Forderungen für ihre Kinder, weil die Wahrscheinlichkeit größer ist, dass ihnen negative Erfahrungen des Umfelds mit dem Bildungssystem bekannt sind (Bourdieu und Passeron 1971). Dieselben Strukturen objektiver Chancen bezüglich des sozialen Aufstiegs durch die Schule bestimmen auch die bildungsbezogenen Einstellungen der Kinder selbst sowie deren Bereitschaft, die dominierenden Normen und Werte des Bildungssystems zu ver-

innerlichen (Bourdieu 2001). Empirisch kann die These Bourdieus von sinkenden Bildungsansprüchen und bildungsbezogenen Werten durch mangelnde objektive Chancen mit einer Untersuchung von Fend (1989) untermauert werden. Lernende gaben in dieser Studie Auskunft über die Bedeutung verschiedener Eigenschaften und Handlungsweisen innerhalb ihrer Schulklasse. Es stellte sich heraus, dass bildungsbezogene Merkmale wie Kompetenz und Schulleistung in Hauptschulklassen zwischen der sechsten und neunten Klasse an Bedeutung verlieren, während an Gymnasien diesen Merkmalen in der neunten Klasse größere Bedeutung zugeschrieben wird als in der sechsten Klasse.

Zusammenfassend können die negativ anmutenden bildungsbezogenen Normen, Werte und Einstellungen der Kinder aus Familien der niedrigeren Schichten bzw. Milieus zwei Mechanismen zugeschrieben werden: Zum einen der Diskrepanz zwischen dem primären Habitus und den schulischen Sinnstrukturen mit den daraus möglichen Passungs- und Anerkennungsdefiziten. Zum anderen der Struktur objektiver Chancen, welche die Bildungsansprüche und -forderungen von Eltern und Kinder mitbestimmen. Es kann also angenommen werden, dass die Passung zwischen den Normen, Werten und Einstellungen der Lernenden und jenen der Schule mit der sozialen Herkunft der Lernenden in Verbindung gebracht werden kann. Dabei ist jedoch nicht davon auszugehen, dass die sozioökonomische Herkunft die Normen, Werte und Einstellungen völlig determiniert. So merkte etwa Bourdieu an, dass es möglich sei, durch die Investition von Zeit den kulturellen Habitus hin zu einer größeren Passung mit der dominierenden Bildungsauffassung zu verändern (Bourdieu 2001). Zudem lässt sich Helspers Theorie bezüglich des primären Habitus und der Passung zu dem jeweiligen Sinnentwurf in der Schule, wie oben gezeigt, zwar unter Hinzunahme von Bourdieu zur Herleitung von Passungsproblemen bei Lernenden mit niedriger sozialer Herkunft anwenden. Es finden sich jedoch Passungs- und Anerkennungsprobleme bei unterschiedlichen primären Habitus mit den jeweiligen Entwürfen des schulisch geforderten Habitus. So finden sich auch Beispiele von Lernenden mit hohem sozialen Status, deren primärer Habitus nicht zum Sinnentwurf der Schule passt (Böhme 2000; Helsper und Kramer 2001). Es mag also durchaus sein, dass bestimmte bildungsbezogene Werte und Normen mit der sozialen Schicht in Zusammenhang stehen, diese Unterschiede sind aber Teil komplexer Exklusionsmechanismen. Es stellt sich daher die Frage, inwieweit die Analyse sozialer Komposition überhaupt zur Erklärung schulischen Lernens beiträgt. Deshalb sollten in quantitativen Studien Normen,

Werte und Einstellungen als eigenständiges Kompositionsmerkmal untersucht werden.

3.4.2 Normative Peer-Effekte

Die Lernenden bringen Normen, Werte und Einstellung in die Lerngruppe ein, die sie durch familiäre Sozialisation sowie im näheren sozialen Umfeld erworben haben. Es wird jedoch angenommen, dass sich die Kinder darüber hinaus in der stetigen Interaktion im Klassenzimmer in ihren Normen, Werten und Einstellungen gegenseitig beeinflussen können. Mit diesem Ansatz erklären beispielsweise Dreeben und Barr (1988) den Abfall der Leistung von Lernenden nach einem Wechsel in eine Schule mit niedrigerem durchschnittlichem sozialem Hintergrund. Die vorherrschenden Normen, Werte und Einstellungen innerhalb der neuen Schule hätten die individuellen Einstellungen der Schulwechsler negativ beeinflusst.

Zur Erklärung dieser Anpassung verweisen Dreeben und Barr (1988) auf die Theorie Meads. Diese Theorie wird der Kritik gerecht, dass Normen, Werte und Einstellungen nicht einfach passiv übernommen werden (wie beispielsweise von Durkheim (1962) oder Parsons (1949) angenommen), vielmehr werden geteilte Normen und Werte durch Artikulation und kontinuierliche Interaktion geschaffen (Eder und Sanford 1986). Eine Ausführung zur Theorie und der Übertragung auf den schulischen Kontext wurde bei Dreeben und Barr (1988) nicht vorgenommen und soll im Folgenden ergänzt werden. Zudem wird die Theorie Meads um entwicklungspsychologische Aspekte und die Kommunikation unter Gleichaltrigen erweitert, da die Peerforschung die besondere Stellung und Einflussnahme von Gleichaltrigen ab einem Alter von ca. 10 Jahren betont (Köhler 2016), welche in der Theorie Meads nur marginal angedeutet ist. Es ist anzunehmen, dass die Kommunikation von Normen und Werten bei Heranwachsenden anders verläuft als bei Erwachsenen (Eder und Sanford 1986). Diesem Umstand wird unter Hinzunahme der Theorie von Piaget (1986) und Youniss (1982) Rechnung getragen, welche die Aushandlungsprozesse unter Kindern und Jugendlichen näher untersucht haben.

Mead sieht die Herausbildung der Identität sowie die Verinnerlichung von Normen und Werten als soziales Phänomen, das heißt begründet durch den Aushandlungsprozess der Person mit gesellschaftlichen Normen und Werten:

„Das [I] ist die Reaktion des Organismus auf die Haltungen anderer; das [Me] ist die organisierte Gruppe von Haltungen anderer, die man selbst einnimmt. Die Haltungen der anderen bilden das organisierte [Me], und man reagiert darauf als ein [I]“ (Mead 1968, S. 216). Das I ist somit der affektive Part. Die Erfahrungen des I gehen in das Me ein, das I reagiert in Rückgriff auf das Me.

Identität ist nach Mead demnach nicht von Geburt an gegeben, sondern entsteht durch die Kommunikation der Person mit der Gesellschaft. Auf eine Äußerung oder Handlung einer Person folgt die Reaktion eines anderen oder mehrerer anderer. Diese Reaktion gibt der Handlung oder Äußerung einen Sinn. Beispielsweise kann durch die Reaktion eines Mitschülers auf die Frage „Hast du die Hausaufgaben gemacht?“ der Handlung, Hausaufgaben zu erledigen, eine positive oder eine negative Bedeutung zugeschrieben werden. Die resultierenden Reaktionen werden von der Person aufgenommen im Me. Durch den stetigen Kommunikationsprozess mit Anderen und auch mit sich selbst ist Identität ein Prozess. Die Einordnung neuer Erfahrungen geschieht unter Rückgriff auf das Me – die bisherigen Erfahrungen (Mead 1968). Somit beeinflussen die ersten Erfahrungen (bspw. im Elternhaus und im näheren sozialen Umfeld) die Reaktion auf neue Erfahrungen.

Die Organisation der Erfahrungen im Me ist von zentraler Bedeutung in der Theorie Meads. Denn alle Erfahrungen müssen im Sinne eines einheitlichen Systems („der Gesamtgesellschaft“) geordnet werden. Verschiedene Erfahrungen können im Konflikt zueinander stehen, beispielsweise vereinnahmt die Gesamtgesellschaft verschiedene Untergruppen wie Parteien oder Verbände mit je eigenen Haltungen. Der Einzelne kann die organisierten Haltungen einer Untergruppe übernehmen, diese müssen jedoch letztlich in Bezug zur übrigen Gesellschaft geordnet werden (Mead 1968). Die Haltung der Gesamtgesellschaft, Bildung einen hohen Wert zuzuschreiben, könnte beispielsweise mit einer negativen Haltung gegenüber Hausaufgaben in Konflikt stehen. Die Reaktion des I – erledigt der/die Lernende die Hausaufgaben oder nicht – ist in diesem Konflikt nicht planbar oder vorhersagbar, jedoch wird eine Reaktion – erledigen oder nicht erledigen – erfolgen. Reaktionen hierauf werden wiederum in das Me aufgenommen und über die Dauer zu einer Lösung des Konfliktes führen: Bspw. Hausaufgaben machen keinen Spaß, aber sie müssen erledigt werden, da Bildung wichtig ist.

Basierend auf dieser Theorie ist Identität geprägt von den verfügbaren Kommunikationspartnern, durch welche Erfahrungen gesammelt werden können; in den ersten Lebensjahren ist dies überwiegend die Familie und das nähere soziale Umfeld. Mit zunehmendem Alter verbringt das Kind mehr Zeit außerhalb der Familie, in Spielgruppen, Vereinen und nicht zuletzt in organisierten Bildungseinrichtungen. Ab dem schulpflichtigen Alter verbringen die Kinder dann einen Großteil ihrer Zeit innerhalb der Schule und besonders innerhalb des Klassenverbandes. Durch die gehäuften Reaktionen durch Mitschüler und Mitschülerinnen ist ein prägender Einfluss auf das Wertesystem des Einzelnen sehr wahrscheinlich.

Offen bleibt in Meads Theorie die Frage, warum der Einfluss der Familie oder des pädagogischen Personals, welche ebenfalls als Interaktionspartner fungieren, zurückgeht und die Jugendlichen untereinander einen stärkeren Einfluss auf die Normen und Werte bekommen. Zur theoretischen Herleitung dieser Annahme soll die Theorie Piagets zur Entwicklung des Regelverständnisses von Kindern sowie die Theorie Youniss' zur Beziehung unter Gleichaltrigen ergänzt werden.

Jean Piaget hat sich in seinem Werk „Das moralische Urteil beim Kinde“ (1986) mit dem Bewusstsein von Regeln und Verhaltensnormen auseinandergesetzt. Ein besonderer Fokus liegt im Gegensatz zu anderen Ansätzen (bspw. Kohlberg 2014; Oser 1981) (Kohlberg 2014; Oser 1981)¹¹ auf den Aushandlungsprozessen von Gruppenregeln bei Kindern, weshalb die an Piagets Werk abgeleiteten Theorien für die Anpassung von Normen und Werten besonders relevant erscheinen und auch häufig in der Peerforschung Anwendung finden (siehe bspw. Köhler 2016; Köhler et al. 2016; Salisch 2016; Steinhoff und Grundmann 2016).

11 Oser beschreibt den interaktiven Prozess von 15-jährigen Schülerinnen und Schülern bei der Diskussion moralischer Dilemmata. Er nimmt dabei vier Stufen der Interaktion an, welche durch die Struktur der Kommunikation und den Einbezug von Normen und Werten geprägt ist. Es geht dabei eher um den kommunikativen Charakter und die Anwendung existierender oder vorgegebener Regeln und Normen. Nicht um die Entwicklung gruppeneigener Normen bei den Heranwachsenden. Ziel ist es, die Heranwachsenden mittels geeigneter Settings in eine höhere Stufe der Interaktion zu bringen. Die Bewertung der Gruppenlösung sowie Änderungen in den Norm- und Regelurteilen spielen keine Rolle.

Piaget identifiziert in seinen Untersuchungen an Kindern unterschiedlichen Alters Entwicklungsstufen im kindlichen Spiel. Die ersten Lebensjahre sind von motorischen und wiederholenden Spielen geprägt. Regeln von Seiten der Eltern werden vom Kind als interessant aufgenommen, aber nicht als verpflichtende Wirklichkeit wahrgenommen. Bezogen auf Mead, kommuniziert das Kind und erhält erste Rückmeldung von Seiten der Eltern. Die Rückmeldungen gehen teilweise in das Me ein, sind aber noch nicht geordnet, um reflektiertes Handeln zu ermöglichen. Mit zunehmendem Alter beginnen die Kinder in Piagets Untersuchungen auferlegte Regeln als verpflichtend zu erleben. In dieser Phase werden die Regeln als von höherer Instanz gegeben und unveränderbar angenommen. Handlungen basieren in dieser Phase vorwiegend auf der Nachahmung der Regeln. In dieser Phase ist davon auszugehen, dass die Kinder überwiegend entsprechend der von den Eltern vermittelten Normen und Werte handeln. Eine bedeutende Veränderung im Regelbewusstsein und im Spiel findet Piaget bei Kindern im Alter von etwa 10 bis 11 Jahren – wenn der Übertritt in die Sekundarschule ansteht (Piaget 1986). War in der vorherigen Phase, in der Regeln als verpflichtend angenommen wurden, das gemeinsame Spiel mehr durch ein nebeneinander Spielen gekennzeichnet und weniger durch kollektive Aktionen, findet Piaget nun ein kooperatives Agieren unter den Kindern. Die Kinder sind in der Lage, Regeln zu reflektieren und nehmen sie als verhandelbar an. Mit Beginn der Sekundarstufe findet sich demnach eine Veränderung in der Denkleistung der Kinder. Die Informationen durch das Gegenüber werden reflektiert und – im Sinne Meads – in einen Gesamtzusammenhang gebracht. Informationen aus verschiedenen Gruppen, wie der Familie und oder der Schulklasse, stehen eventuell im Konflikt zueinander und müssen innerhalb des Me verbunden und geordnet werden. Dies öffnet den Raum, das überwiegend durch die Familie geprägte Wertesystem in Interaktion mit den Klassenkameraden zu erweitern und umzustrukturieren.

Warum gerade die Reaktionen der Mitschüler und Mitschülerinnen große Relevanz erhalten, und weniger jene von Erwachsenen, kann durch die Theorie von Youniss zur Beziehung von Interaktionspartnern erklärt werden (Köhler 2016; Salisch 2016). Youniss (1982) unterscheidet zwei Arten der reziproken Interaktion: die komplementäre und die symmetrische. In der komplementären Reziprozität ist die Handlung des einen Beteiligten kontrolliert oder gelenkt durch den Interaktionsbeitrag des anderen. Die Interaktion zwischen Erwachsenem und Kind findet üblicherweise in dieser Form statt. Der Erkenntnisgewinn ist meist einseitig. In der symmetrischen Reziprozität hingegen tragen zwei (oder

mehr) Personen gleichartige Handlungen zur Interaktion bei. Die Kontrolle sowie der interpersonale Einfluss sind wechselseitig und es kommt zum Erwerb von Wissen auf beiden Seiten. In dieser Form der Interaktion werden Informationen, Ansichten und Handlungen kritisch reflektiert. Auf diese Weise entwickeln Peers gemeinsame Regeln und Vorstellungen, wie etwas zu funktionieren hat (Siegler et al. 2016).

Bezogen auf Meads Theorie könnten die Interaktionen unter den Gleichaltrigen die Ordnung neuer Reaktionen, welche eventuell im Widerspruch zur bisherigen Struktur des Me stehen, erleichtern und somit die Koexistenz und Veränderung von Normen, Werten und Einstellungen ermöglichen. Durch den Zusammenschluss innerhalb der Schulklasse sind die Heranwachsenden in besonderem Maße gezwungen einen gemeinsamen Konsens zu finden. Anders verhält es sich beispielsweise auf öffentlichen Plätzen, bei denen sich Individuen und Gruppen mit oppositionellen Normen, Werten und Einstellungen aus dem Weg gehen können (Krappmann 2004). Gerade dieses regelmäßige und vorhersehbare Aufeinandertreffen innerhalb der Schulklasse intensiviert die Herausforderung, gemeinsame Normen und Werte für das gemeinsame Lernen innerhalb der Lerngruppe auszuhandeln. Je nach den herrschenden Normen bezüglich der gegenüber der Schule einzunehmenden Haltung kann so beispielsweise persönliches Engagement im Unterricht als Verletzung der herrschenden Normen, Werte und Einstellungen gelten (Breidenstein und Kelle 2002).

Aushandlungsprozesse sind nicht zwingend auf die gesamte Klasse bezogen. Stattdessen wird in mehreren Studien beobachtet, dass sich Subgruppen innerhalb einer Klasse bilden können. Die Schulklasse ist damit nicht nur ein formales System, das extern organisiert ist. Sie kann auch als informales System gelten, in dem Interaktionen zwischen Lernenden durch Sympathie und Freundschaftsbeziehungen beeinflusst werden (Krappmann 2004).

Solche Formationen von Subgruppen finden bspw. Adler und Adler (2003) in ihrer ethnographischen Studie von Lernenden über mehrere Schuljahre hinweg. Sie beschreiben, dass sich ab der siebten Klasse Subgruppen formieren und ab ca. der achten Klasse feste Gruppenstrukturen vorherrschen. Ähnlich wie bei der Entwicklung der Aushandlungsprozesse scheinen stabile soziale Gruppen erst mit zunehmendem Alter bedeutend zu werden. Die soziale Herkunft der Heranwachsenden scheint bei der Bildung von Gruppen ein nicht zwingendes, aber durchaus relevantes Merkmal zu sein (Adler und Adler 2003). Dies konnte

auch in anderen Studien beobachtet werden (bspw. Eder und Hörl 2010; Fend und Specht 1977; Helsper 1989; Willis 1982). Es wird vermutet, dass mit den ähnlichen Lebensumständen und -erfahrungen eine positive Bestätigung innerhalb der Gruppe verbunden ist (Bierhoff 2002).

Diese positive Bestätigung gibt dem Individuum Sicherheit, welches bestrebt ist, die eigene Meinung und das Verhalten zu validieren. Da für Meinungen und Einstellungen meist objektive Kriterien fehlen, wird auf soziale Vergleiche zurückgegriffen (Festinger 1950). Eine abweichende Meinung führt zu Unsicherheit, weshalb Personen dazu tendieren, sich Gruppen anzuschließen, in denen allgemein ähnliche Meinungen und Fertigkeiten vorherrschen. Im Umkehrschluss werden Kontakte mit stark divergenten Personen vermieden und diese nicht als vergleichbar wahrgenommen.

Durch solche Gruppierungsprozesse können mehrere Meinungen parallel in einer Gesellschaft existieren. Bezogen auf die Schulklasse können sich Subgruppen mit divergenten schulbezogenen Normen, Werten und Einstellungen herausbilden. Solche Unterschiede konnten beispielsweise in Klassen mit soziokulturell sehr heterogenen Lernenden gefunden werden. Eine solche Komposition führt jedoch nicht per se zu divergenten Subgruppen. Stattdessen ist der Hang nach sozialem Vergleich und übereinstimmenden Meinungen, sowie damit zusammenhängend die Tendenz, Subgruppen zu bilden und die Konformität zu erhöhen, von äußeren Faktoren abhängig. Je relevanter das betreffende Thema ist und je attraktiver die Gruppe ist, desto eher finden die beschriebenen Gruppierungsprozesse statt (Festinger 1950). Kommt beispielsweise einer bestimmten Schülergruppe mit geringer Passung zum geforderten schulischen Habitus eine vergleichsweise geringe Anerkennung zu, können die Schule sowie schulkonforme Schülergruppen als Bedrohung angesehen werden. Eine fehlende Anerkennung bzw. Abwertung von Seiten der Schule kann dann das Bedürfnis nach Sicherheit und sozialer Bestätigung innerhalb der eigenen Gruppe verstärken und den Druck nach Konformität innerhalb der Gruppe erhöhen (Festinger 1950). Vor allem in Abgrenzung zu anderen Gruppen können dann gruppenspezifische Verhaltensweisen verstärkt werden, um die Zugehörigkeit zur eigenen Gruppe zu signalisieren (Austan-Smith und Fyer 2005). Dies geschieht beispielsweise über eine verstärkte Kommunikation innerhalb der Gruppe mit den Personen, die von der mittleren Gruppenmeinung abweichen mit dem Ziel, eine stärkere Homogenität innerhalb der Gruppe herzustellen. Werden die Divergenzen einzelner Gruppenmitglieder zur mittleren Grup-

penmeinung zu groß, erfolgt hingegen der Ausschluss aus der Gruppe, beispielsweise indem nicht mehr mit diesen Personen kommuniziert wird. Auf Seiten des Individuums ist eine Anpassung der Meinung zu erwarten, vor allem wenn die anderen Gruppenmitglieder sehr homogen bezüglich der Meinung sind, es keine passendere Gruppe für die eigene Meinung gibt und die Attraktivität der Gruppe bzw. das Zugehörigkeitsbedürfnis hoch ist (Festinger 1950).

Für den schulischen Kontext lässt sich daraus folgern, dass es in Klassen mit stark heterogener Schülerschaft in Bezug auf Normen, Werte und Einstellungen zur Bildung divergenter Subgruppen kommen kann, wobei die Bedeutung der Subgruppe und der Druck nach Konformität von äußeren Faktoren abhängen. Diese näher zu untersuchen, würde das Verständnis für Peer-Interaktionen und Effekte der Komposition erhöhen. In Klassen mit einer eher homogenen Schülerschaft haben einzelne divergente Lernende weniger Möglichkeiten, die Vergleichsgruppe zu wechseln. Es ist anzunehmen, dass die Tendenz groß ist, die eigene Meinung entsprechend der Gesamtgruppe anzupassen, um wahrgenommene Unsicherheiten zu reduzieren.

3.4.3 Instrumentelle Peer-Effekte

Werden keine interaktiven Anpassungsprozesse in den Blick genommen, sondern die Summe der vorhandenen Eigenschaften und Verhaltensweisen der Lernenden einer Klasse, wird auch von instrumentellen Peer-Effekten gesprochen (Harris 2010). Die Eigenschaften und Verhaltensweisen der Lernenden gelten hier zum einen als geteilte Ressource und zum anderen als Ausgangspunkt für das Lernumfeld. Die Idee ist, dass bei einem hohen Anteil von Lernenden mit geringer Aufmerksamkeit, mangelnder Disziplin und oppositionellem Verhalten sich die Lernenden gegenseitig Ressourcen wegnehmen. In Klassen mit Disziplinschwierigkeiten und somit häufigen Störungen reduziert sich beispielsweise die Zeit, welche die Lernenden für Instruktion und Interaktion mit den Lehrenden zur Verfügung haben, da die Lehrenden Zeit investieren müssen, um Unterrichtsstörungen zu unterbinden. Unterrichtsstörungen mindern somit die verfügbare Instruktionszeit und Interaktionszeit (Manski 2000). Dies kann mit dem Unterrichtmerkmal „Disziplin“ bzw. „disziplinäres Klassenklima“ in Verbindung gebracht werden.

Die Leistungen und Fähigkeiten der Lernenden können ebenfalls als Ressource betrachtet werden. In Klassen mit einem großen Anteil an Lernenden mit hoher kognitiver Fähigkeit können sich die Lernenden gegenseitig gut unterstützen, indem sie ihr Wissen und Können im Sinne einer Ressource miteinander teilen und kognitiv anregende Unterrichtsgespräche führen. In Diskussionen können hier verschiedene Meinungen gegenübergestellt werden (Rjosk 2015). Laut Piagets (1977) Konzept des Disäquilibriums bewirkt die Konfrontation mit neuer Information, vor allem wenn sie nicht zur eigenen Wissenskonstruktion passt, einen kognitiven Konflikt. Die intendierte Abwendung dieses Zustands ruft durch Prozesse der Assimilation und Akkommodation der neuen Wissensinhalte eine kognitive Weiterentwicklung hervor.

Auch ein hoher Anteil an Lernenden mit Deutsch als Zweit- bzw. Drittsprache ist im Blickwinkel der Ressourcen zu sehen (Esser 2006). Hier werden jedoch meist negative Effekte angenommen. Argumentiert wird, dass die Häufigkeit, die Unterrichtssprache zu sprechen und zu hören, eine Ressource ist, die sich mit steigendem Anteil an Lernenden mit gleicher unterrichtsferner Verkehrssprache reduziert (Rjosk et al. 2014). Dadurch hätten die Heranwachsenden weniger sprachliche Rollenmodelle der Unterrichtssprache zur Verfügung und würden den korrekten Sprachgebrauch der Unterrichtssprache weniger üben (Esser 2009; Hopf 2005). Hierdurch sind primär die sprachlichen Schülerleistungen betroffen, in der Konsequenz aber auch die Leistungen in anderen Fächern (van Ewijk und Slegers 2010). Dem stehen Überlegungen entgegen, dass Lernende, welche die gleiche Sprache zusätzlich zur Instruktionssprache sprechen, sich gegenseitig unterstützen können, indem sie sich das Lernmaterial in ihrer Herkunftssprache erklären (Maestri 2011). Untersuchungen zur bilingualen Interaktion in einer anderen Sprache als der Unterrichtssprache konnten bspw. zeigen, dass die Interaktionen zur Rückversicherung genutzt werden (Swain und Lapkin 2000) sowie unbekannte Wörter der Unterrichtssprache übersetzt oder erklärt werden (Hopewell 2011; Klingner und Vaughn 2000). Zudem wird argumentiert, dass der Gebrauch der Herkunftssprache neben den direkt erzielten Kompetenzzuwächsen in der Herkunftssprache auch zu den erwünschten Fortschritten im Deutschen führt (Gogolin, Neumann & Roth, 2005; Neuner, 2003). Diese Position der Verbundenheit sprachlicher Kompetenzen in mehreren Sprachen geht auf Cummins (2006) Interdependenzhypothese zurück. Diese postuliert, dass kognitiv akademische Sprachfähigkeiten, wie bspw. phonologische Bewusstheit, Lesestrategien, Rechtschreibung, zwischen den Sprachen von Individuen transferiert werden können. Somit kann der

Anteil Lernender mit Migrationshintergrund auch im positiven Sinne als unterstützende Ressource betrachtet werden.

Zusätzlich zu den Ressourcen wird die Gesamtheit der Eigenschaften und Verhaltensweisen der Lernenden auch mit Blick auf das Lernumfeld diskutiert (Stäbler 2017). Ein wichtiger Faktor des Lernumfeldes ist das Klassenklima. Das Klassenklima nimmt in der Unterrichtsforschung einen besonderen Stellenwert ein, da ein positiver Zusammenhang zwischen förderlichem Klassenklima und der Entwicklung der Lernenden mehrfach belegt werden konnte. Verschiedene Studien belegen bspw. einen positiven Zusammenhang zwischen dem Klima und der Schulleistung (einen Überblick gibt Eder 2010). Zudem finden sich Hinweise, dass ein positives Klassenklima förderlich auf psychosoziale Merkmale wie Selbstwert, allgemeines Wohlbefinden, Lernmotivation, Lernzufriedenheit und Lernengagement wirkt (Cohen et al. 2009; Eder 2010; Scharenberg und Bos 2014; Thapa et al. 2013). Die Grundannahmen für den Zusammenhang zwischen der Klassenkomposition und dem Klassenklima beziehen sich auf oben beschriebene individuelle Merkmale, welche bestimmten Schülergruppen zugeschrieben werden (siehe Abschnitt 3.4.1). Individuelle Schülerinnen und Schüler mit bestimmten Hintergrundmerkmalen (wie kein Zuwanderungshintergrund, privilegierte soziale Herkunft, hohes Leistungspotenzial) sollen demnach eine hohe Lernmotivation sowie lernförderliche Wertvorstellungen, Überzeugungen und Verhaltensweisen haben. Wenn viele solcher Lernender in einer Klasse gemeinsam beschult werden, können sie gemeinsam eine normative „Kultur des Erfolgs“ schaffen. Die Kehrseite kann eine Peerkultur darstellen, in der ein gemeinsames Gefühl von Mangel und Selbstzweifel herrscht, kollektives Stigmatisierungserleben, Gefühl einer Entfremdung von der Schule sowie Akzeptanz abweichenden Verhaltens, wie Störverhalten im Unterricht oder Absentismus (Baumert et al. 2006b; Dreeben und Barr 1988; Harker und Tymms 2004; Rumberger und Palardy 2005)

Die Unterrichtsqualität ist ein weiteres wichtiges Merkmal zur Beschreibung des Lernumfeldes. Hier gibt es für die soziale Zusammensetzung Hinweise darauf, dass Effekte des durchschnittlichen sozioökonomischen Status der Schule oder Klasse durch Merkmale der Unterrichtsqualität vermittelt werden (Rumberger und Palardy 2005). So zeigten sich positive Zusammenhänge zwischen dem durchschnittlichen sozialen Status und Merkmalen der Unterrichtsqualität, deren Kontrolle soziale Kompositionseffekte auf Schülerleistung reduzierte. In der Mehrzahl der genannten Studien wurde jedoch die Schulebene

betrachtet, die soziale und zuwanderungsbezogene Zusammensetzung nicht gemeinsam untersucht bzw. weitere leistungsbezogene Merkmale der Zusammensetzung außer Acht gelassen. Somit stellt die Untersuchung der vermittelnden Rolle der Unterrichtsqualität für Effekte der sozialen und zuwanderungsbezogenen Klassenzusammensetzung auf Schülerleistung weiterhin ein Forschungsdesiderat dar (Rjosk 2015).

3.4.4 Big-Fish-Little-Pond-Effekt

Eine etwas andere Art des Peer-Effektes stellt der Big-Fish-Little-Pond-Effekt (BFLPE) dar, welcher sich auf den Zusammenhang der Klassenkomposition mit psychosozialen Merkmalen bezieht (Marsh 1987). Im BFLPE-Modell wird angenommen, dass Lernende ihre eigenen akademischen Fähigkeiten mit denen ihrer Klassenkameraden vergleichen und diese sozialen Vergleichsinformationen als eine Grundlage für die Bildung ihres eigenen akademischen Selbstkonzepts und für andere Selbstbeurteilungen verwenden. Ein negativer Effekt bedeutet, dass ein Schüler bzw. eine Schülerin ein niedrigeres akademisches Selbstkonzept hat, wenn er bzw. sie sich mit leistungsstärkeren Lernenden vergleicht. Derselbe Schüler bzw. dieselbe Schülerin hat hingegen ein positiveres akademisches Selbstkonzept beim Vergleich mit weniger leistungsstarken Lernenden (Marsh et al. 2008).

Sowohl im internationalen Kontext wie auch für Deutschland konnte dieser Effekt gezeigt werden (Becker und Neumann 2018; Köller 2004; Köller und Baumert 2001; Marsh 2005; Marsh et al. 2008; Stäbler et al. 2017). Kulturübergreifend wird er begründet über fehlende objektive Kriterien für die eigene Fähigkeitsbeurteilung. Stattdessen wird ein Vergleich mit den kognitiven und nichtkognitiven Outcomes der Bezugsgruppe vorgenommen, welche im schulischen Kontext üblicherweise die Klasse darstellt. Vergleicht sich ein Schüler bzw. eine Schülerin stets mit Mitschülerinnen und Mitschülern in einer Klasse mit einem verhältnismäßig niedrigen Leistungsniveau (little pond) schätzt sich dieser Schüler bzw. diese Schülerin als guter Schüler bzw. gute Schülerin ein (big fish). In einer Klasse mit hohem Leistungsniveau (big pond) würde sich der gleiche Schüler bzw. die gleiche Schülerin hingegen als schwach einschätzen (Zurbriggen 2016).

Obwohl der BFLPE über verschiedene Settings hinweg repliziert werden konnte, ist bislang wenig erforscht, wie stabil das akademische Selbstkonzept bei wechselndem Kontext ist. Bislang scheinen nur fünf Studien zu existieren, die sich dieser Frage widmen (Becker und Neumann 2018). Die Befundlage ist in diesen Studien nicht eindeutig, was aus den abweichenden Zielpopulationen resultieren könnte. Drei der Studien untersuchen die Stabilität des akademischen Selbstkonzeptes am Übergang von der Sekundarschule in anschließende Bildungs- und Arbeitsmöglichkeiten (Marsh 1991; Marsh et al. 2016; Marsh und O'Mara 2010). In allen drei Studien zeigte sich das akademische Selbstkonzept als relativ beständig. Zwei Studien, welche das akademische Selbstkonzept am Übergang von der Primar- in die Sekundarstufe untersuchen, finden hingegen abweichende Befunde (Becker und Neumann 2016, 2018). Becker und Neumann (2018) führen dies unter anderem auf die geänderte Zusammensetzung der Lerngruppe durch das gegliederte Bildungssystem zurück. Insgesamt fehlt es jedoch an weiteren Studien um sichere Aussagen über die Stabilität des Selbstkonzeptes in verschiedenen Kontexten und über verschiedene Domänen hinweg treffen zu können.

Auch für andere selbstberichtete psychosoziale Merkmale konnten Zusammenhänge mit der Referenzgruppe im Sinne eines BFLPE gefunden werden. So wurden in einzelnen Studien negative Effekte der mittleren Schulleistung auf das Selbstvertrauen und die Kontrollerwartung von Sekundarschülerinnen und -schülern (Marsh et al. 2008) sowie auf das fachspezifische Interesse von Lernenden der Sekundarstufe gefunden (Köller et al. 2000). Für die mittlere Klassenleistung ergaben sich zudem negative Effekte auf die Berufsaspirationen von Sekundarschülerinnen und -schülern (Zurbriggen 2016). Neben der leistungsbezogenen Zusammensetzung scheint auch die soziale Zusammensetzung der Lerngruppe mit der Ausprägung psychosozialer Merkmale zusammenzuhängen. So fand Crosnoe (2009) Hinweise darauf, dass Lernende an Schulen mit niedrigem sozioökonomischem Status sich weniger sozial isoliert fühlten als Lernende an Schulen mit höherem mittlerem sozialem Status. Für die Einschätzung von Lernenden bezüglich der Unterstützung und Förderung durch die Lehrperson (Schüler-Lehrer-Beziehung) gibt es ebenfalls Hinweise auf einen Effekt der Bezugsgruppe. So finden Vieluf et al. (2015) einen Zusammenhang zwischen der schulischen sozialen Komposition und der selbsteingeschätzten Schüler-Lehrer-Beziehung in gegliederten Schulsystemen. Hier berichten Lernende in Schulen mit einem niedrigeren mittleren sozioökonomischen Status von einer qualitativ höheren Schüler-Lehrer-Beziehung als Lernende in Schu-

len mit höherem mittlerem sozioökonomischem Status. Vieluf et al. (2015) erklären dies unter anderem damit, dass in Schulen mit geringerem mittlerem sozialem Status ein schlechteres soziales Klima herrschen könnte. Lernende dieser Schulen könnten somit geringere Erwartungen an die Beziehung zu ihren Lehrenden entwickeln. Mit Bezug zu ihrer Referenzgruppe könnten sie dann die Schüler-Lehrer-Beziehung positiver einschätzen als vergleichbare Beziehungen in Schulen mit positiverem Klima.

Betrachtet man also nicht nur die Leistungsentwicklung der Lernenden, sondern auch psychosoziale Variablen, so zeigt sich, dass eine Klasse oder Lerngruppe mit eher niedrigeren mittleren Fähigkeit bzw. Leistung sowie mit niedrigem mittlerem sozialem Status auch positive Effekte für die Entwicklung des Einzelnen haben kann. Demnach kann eine bestimmte Zusammensetzung einer Schülerschaft je nach betrachteter abhängiger Variable unterschiedliche, jedoch gleichzeitig wirkende Einflüsse haben. Aufgrund der bisherigen Befundlage zu Kompositionseffekten ist allerdings davon auszugehen, dass für die Leistungsentwicklung mit negativen Effekten zu rechnen ist, wenn die Schülerschaft im Mittel eher niedrige Leistungen zeigt sowie einen niedrigen mittleren sozialen Status aufweist.

3.4.5 Erwartungen der Lehrenden

Neben den Peer-Effekten als Wirkmechanismen der Komposition werden auch Prozesse auf Seiten der Lehrenden als Mediatoren für die Wirkung der Zusammensetzung einer Schülerschaft auf die individuellen Schülerleistungen angenommen. Annahmen zu Lehrererwartungen werden vor allem in der psychologischen Forschung diskutiert und beziehen sich darauf, dass Lehrende von Lernenden mit Zuwanderungshintergrund oder mit niedrigerem sozioökonomischem Status geringere Leistungen erwarten als von Heranwachsenden ohne Zuwanderungshintergrund oder mit einem sozioökonomisch privilegierten Status. Daran anknüpfend wird davon ausgegangen, dass Lehrende den Erwartungen entsprechend wiederum ihr Verhalten gegenüber den Lernenden sowie das unterrichtliche Handeln anpassen, was wiederum die schulischen Leistungen beeinträchtigen kann (Dreeben und Barr 1988; Harker und Tymms 2004; Harris 2010).

Erste Hinweise auf den Zusammenhang zwischen Erwartungen von Lehrenden und den Entwicklungsmöglichkeiten ihrer Schülerinnen und Schüler liefert die vieldiskutierte Studie von Rosenthal und Jacobson (1968) zum Pygmalion-Effekt. Inzwischen sind Zusammenhänge zwischen der Lehrererwartung und der Schülerleistung mehrfach repliziert worden und finden allgemeine Akzeptanz (Rubie-Davies 2009). Dass Lehrende ihre Erwartungen abhängig von der Gruppenzusammensetzung anpassen, konnte ebenfalls gezeigt werden. Beispielsweise finden Dreeben und Barr (1988) vermittelnde Effekte der Leistungskomposition auf die Leseleistung über die Lehrererwartung. Zu ähnlichen Befunden für die soziale Komposition kamen Rumberger und Palardy (2005), die einen vermittelnden Effekt der sozialen Komposition einer Schule auf Schulleistungen über die Erwartungen der Lehrenden und das vorherrschende akademische Klima beobachteten.

Es wird vermutet, dass der Pygmalion-Effekt durch bestimmte Stereotype erzeugt wird, welche Lehrende dazu verleiten, sich gegenüber den Lernenden in einer bestimmten Weise zu verhalten. Schuchart und Dunkake (2014) können solche schulbezogenen Stereotype, welche mit der sozialen Herkunft sowie mit der Leistung von Lernenden verbunden sind, bereits bei angehenden Lehrkräften finden. Lernende aus niedrigen Sozialgruppen werden ähnliche, vorwiegend negative Eigenschaften zugeschrieben wie leistungsschwachen Lernenden. Mit hoher Übereinstimmung wird beiden Schülergruppen – Leistungsschwachen sowie Lernenden aus den unteren Sozialgruppen – zugeschrieben, sie seien uninteressiert, unkonzentriert sowie undiszipliniert. Leistungsstarke Lernende sowie Lernende aus höheren sozialen Gruppen werden hingegen überwiegend mit positiven Eigenschaften besetzt, bspw. seien sie motiviert, strukturiert und sozial kompetent.

Entsprechend den Erwartungen passen die Lehrenden schließlich ihr Verhalten gegenüber den Lernenden sowie das unterrichtliche Handeln an. Beispielsweise geben Lehrende den Lernenden bei hohen Leistungserwartungen mehr Zeit zum Nachdenken, loben häufiger für Erfolge, kritisieren weniger und interagieren häufiger mit leistungsstarken Lernenden (Brophy 1983; Brophy und Good 1974; Cooper 1979; Cooper und Good 1983). Für das unterrichtliche Handeln konnte gezeigt werden, dass Lehrende mit hohen Erwartungen gegenüber der Lehrgruppe eher unterstützende Ansätze verfolgen, bei niedrigen Erwartungen ist der Unterricht hingegen stärker lehrerzentriert (Rubie-Davies 2009): Bei Lehrenden mit hoher Leistungserwartung findet Unterricht häufiger in leis-

tungsgemischten Kleingruppen statt, die Lehrenden erteilen häufiger lernbezogenes Feedback und stellen anspruchsvollere Fragen. In Klassen von Lehrenden mit geringer Leistungserwartung haben Lernende hingegen wenig Möglichkeiten mit Peers anderer Fähigkeitsniveaus zusammenzuarbeiten, Fragen der Lehrenden sind überwiegend geschlossen formuliert und auf bekanntes Wissen bezogen, es werden wenig Feedback und kaum individuelle Lernziele gegeben und Lehrende wenden vergleichsweise viel Zeit für unterrichtsstrukturierende Anweisungen auf.

Für die Schülerinnen und Schüler einer Lerngruppe können sich demnach entsprechend der leistungsbezogenen und sozialen Zusammensetzung und den damit verbundenen Erwartungen der Lehrperson unterschiedliche Unterrichtsangebote und damit verbundene Lerngelegenheiten ergeben. Es gilt jedoch zu beachten, dass die zitierten Studien überwiegend aus Ländern mit integrierten Bildungssystemen stammen. In institutionell gegliederten Bildungssystemen wie dem deutschen ist es hingegen schwierig, Unterrichtsmerkmale, die eine Reaktion auf die Zusammensetzung der Schülerschaft darstellen, von den anderen systembedingten Unterschieden im Unterricht zu trennen (Gamoran, Nystrand, Berends und LePore, 1995). Die Studie von van Ophuysen und Wendt (2010), welche gezielt Einflüsse von Kompositionsmerkmalen und Unterrichtsmerkmalen untersucht, stellt daher in Deutschland eine Besonderheit dar. Sie untersuchen anhand der Daten des DFG-Projektes „Grundschulübergang“ für Dortmunder Schulen die Bedeutung von Kompositionsmerkmalen in Abgrenzung zu Unterrichtsmerkmalen. Die Forscher schlussfolgern aus ihren Daten, dass Unterrichtsmerkmale unter Berücksichtigung der Schularten eine größere Erklärungskraft für die Leistungsentwicklung der Lernenden haben als Kompositionsmerkmale. Allerdings wurden Zusammenhänge zwischen der Klassenzusammensetzung und Prozessmerkmalen des Unterrichts nicht untersucht. Weshalb hier weiterhin ein großer Forschungsbedarf besteht. Für einzelne Prozessmerkmale des Unterrichts finden van Ophuysen und Wendt (2010) jedoch bedeutende Zusammenhänge mit der Schulart.

3.5 Lehren und Lernen vergleichend zwischen den Schularten – der Forschungsstand

Dieser Abschnitt fasst den Forschungsstand zu Schulartvergleichen im Lehren und Lernen zusammen. Lehren und Lernen meint hier, dass nicht nur Prozessmerkmale zur Beschreibung des Unterrichts einbezogen werden, sondern auch die Komponenten des Aushandlungsprozesses (siehe Abbildung 3). Schulartunterschiede auf Systemebene in der Lehrerbildung, dem Curriculum sowie den Selektions- und Segregationsprozessen inklusive damit einhergehender Kompositionseffekte wurden bereits in den vorangegangenen Abschnitten theoretisch und empirisch behandelt. Hier geht es nun konkret um Unterschiede im Lehren und Lernen an den Schulen der verschiedenen Schularten.

Pionierarbeiten, welche Unterschiede zwischen Schularten im Lehren und Lernen betonen, entstanden bereits im Zuge des Gesamtschulversuchs und den anschließenden Schulstrukturdebatten in den 1970er und 1980er Jahren. Vor allem Fend und seine Kollegen stellten umfassende Analysen von Schul-, Unterrichts- und Individualmerkmalen an mit dem Ziel, die politischen Diskussionen von Fürsprechern und Gegnern der Gesamtschulen auf eine wissenschaftliche Basis zu stellen. Die Analysen verglichen noch nicht direkt die damaligen Schularten, sondern stellten das dreigliedrige System (Gymnasium, Realschule, Hauptschule) dem Gesamtschulsystem (integrierte und kooperierende Gesamtschule) gegenüber. Exkurse der Forscher bieten aber auch differenziertere Einblicke in die einzelnen Schularten (Fend 1976a, 1976b, 1976c; Fend 1982; Fend und Klaghofer 1980; Fend et al. 1976; Fend und Specht 1977).

Aus den Systemvergleichen sowie den Exkursen lassen sich bereits wichtige Erkenntnisse über unterschiedliche Ausgangsbedingungen an den Schularten im Hinblick auf die Zusammensetzung der Schülergruppe und die Einstellungen der Lehrenden folgern. Beispielsweise findet Fend Unterschiede in der Leistungsangst, dem Wohlbefinden und der Schulfreude zwischen dem Gesamtschulsystem und dem traditionellen System, mit günstigeren Konstellationen in den Gesamtschulen (Fend 1982). Damit scheinen negative Konsequenzen normativer Peer-Effekte (siehe Abschnitt 3.4.2) an Gesamtschulen geringer auszufallen. Auch auf Seiten des Lehrerhandelns finden sich bedeutende Unterschiede: Im Gesamtschulsystem ist die konstruktive Unterstützung höher bei gleichzeitig geringerem Anforderungsdruck. Die Lehrenden verbringen weni-

ger Zeit mit Unterrichtsvorbereitung und setzen weniger Veranschaulichungsmaterial im Unterricht ein (Fend 1982).

Aus diesen Ergebnissen können systematische Unterschiede in den Einstellungen der Lernenden sowie in den Einstellungen und dem Handeln der Lehrenden gefolgert werden, wie sie auch in aktuelleren Studien gefunden wurden (siehe Abschnitt 3.5.1 und Abschnitt 3.5.2). Trotzdem zieht Fend in seinen Gesamtschulstudien das Fazit, die Varianz zwischen Schulen innerhalb der Schulsysteme sei bei einigen Merkmalen bedeutend größer als die Variation zwischen den Schulsystemen, weshalb er in seinen folgenden Arbeiten die Bedeutung der Einzelschule betont. Die hohe Varianz zwischen den Schulen innerhalb des traditionellen Systems verwundert nicht, wenn man berücksichtigt, dass Fend und Kollegen den Gesamtschulen einen Komplex von Hauptschulen, Realschulen und Gymnasien gegenüber gestellt haben: „Dabei haben wir in der Regel zwei Hauptschulen, eine Realschule und ein Gymnasium herangezogen und als äquivalent zu einer Gesamtschule definiert“ (Fend 1982, S. 20). Somit werden die Schularten des traditionellen Schulsystems nicht getrennt analysiert, was bei den großen zu erwartenden Differenzen zwischen diesen Schularten, besonders zwischen Gymnasium und Hauptschule, als kritisch zu sehen ist.

Neben den ideologischen Auseinandersetzungen der 1970er Jahre als Grund, die Schularten als traditionelles Bildungssystem und Gesamtschulsystem zusammenzufassen und gegenüberzustellen, könnte auch die geringe Anzahl an Schulen in der Stichprobe ein Hinderungsgrund für systematische Schulartvergleiche gewesen sein. Umfassende, alle Schularten und Bundesländer umfassende Analysen sind erst mit den großangelegten Leistungs- und Vergleichsstudien seit Ende der 1990er Jahre möglich (Baumert, Lehmann, Lehrke, Schmitz, Clausen, Hosenfeld, Köller 1997). Die größten Leistungsvergleichsstudien im internationalen Bereich sind „Programme for International Student Assessment“ (PISA), „Trends in International Mathematics and Science Study“ (TIMSS) und „Progress in International Reading Literacy Study“ (PIRLS), welche durch ihre nationalen Erweiterungen auch bedeutende Informationen zu den Prozessen des Unterrichts in Deutschland liefern. Nationale Studien wie „Deutsch Englisch Schülerleistungen International“ (DESI), „Bildungsverläufe und psychosoziale Entwicklung im Jugend- und jungen Erwachsenenalter“ (BIJU), „Kompetenzen und Einstellungen von Schülerinnen und Schülern“ (KESS) oder der IQB Bildungstrend ergänzen die internationalen Studien. Qualitative Studien können nur selten Vergleiche zwischen Schularten anstel-

len, da Verallgemeinerungen mit den meist geringen Fallzahlen problematisch sind. Doch auch in diesem Forschungsbereich liefern einzelne Studien detaillierte Erkenntnisse über das Lehren und Lernen in den einzelnen Schularten (bspw. Bohl et al. 2013). Die Erkenntnisse aus den genannten Studien zu Schulartunterschieden werden im Folgenden systematisch zusammengefasst und soweit vorhanden um Erkenntnisse aus kleineren empirischen Forschungsprojekten sowie qualitativen Erkenntnissen ergänzt.

3.5.1 Lernende: Leistung, psychosoziale Merkmale und Differenzielle Lern- und Entwicklungsmilieus

Häufigstes Merkmal, welches in den Leistungsvergleichsstudien vergleichend zwischen den Schularten betrachtet wird, ist – wie die Bezeichnung bereits verrät – die Testleistung der Lernenden in verschiedenen Domänen. Trotz der mannigfaltigen Forschungsergebnisse zur Schülerleistung wird nur sehr knapp auf diese Befundlage eingegangen. Aufgrund der hohen Übereinstimmung in den Ergebnissen erscheint dies auch gerechtfertigt.

Vielfach rezipiert ist die Erkenntnis, dass zwischen den Schularten der Sekundarstufe I bedeutende Unterschiede in den mittleren Leistungen vorliegen. Lernende in Hauptschulen verzeichnen die geringsten mittleren Testleistungen, gefolgt von Lernenden an Schulen mit mehreren Bildungsgängen, integrierten Gesamtschulen und Realschulen, wobei die mittleren Schülerleistungen dieser Schularten relativ nahe beieinander liegen. Die – meist mit deutlichem Abstand – höchsten Testleistungen erzielen im Durchschnitt Lernende an Gymnasien. Dieses wiederkehrende Muster findet sich über die verschiedenen Studien und Fächer hinweg, bspw. in

- Mathematik (z. B. Baumert et al. 2004; Baumert, Lehmann, Lehrke, Schmitz, Clausen, Hosenfeld, Köller et al. 1997; Guill et al. 2010; Kunter et al. 2005; Schiepe-Tiska et al. 2013; Wendt et al. 2010)
- Deutsch (z. B. Gailberger und Willenberg 2008; Naumann et al. 2013; Thomé und Eichler 2008; Wendt et al. 2010; Willenberg 2008)
- Englisch (z. B. Harsch et al. 2008; Max-Planck-Institut für Bildungsforschung 1996; Nikolova und Ivanov 2010)

- Naturwissenschaftliche Fächer (bspw. Baumert, Lehmann, Lehrke, Schmitz, Clausen, Hosenfeld, Köller 1997; Ivanov und Nikolova 2010; Max-Planck-Institut für Bildungsforschung 1996; Prenzel et al. 2004)

Vor dem Hintergrund der leistungsbezogenen Selektion im deutschen Schulsystem erscheinen diese Ergebnisse plausibel. Entgegen den Erwartungen gibt es jedoch große Überschneidungen in den Leistungsverteilungen zwischen den Schulformen (Baumert et al. 2006b; Sälzer, Prenzel et al. 2013). So erreichen bspw. 40 % der Realschüler in Mathematik das durchschnittliche Niveau von Gymnasiasten, 25 % liegen sogar darüber. In den naturwissenschaftlichen Leistungen ist die Verteilung ähnlich (Baumert et al. 2006b; Baumert, Lehmann, Lehrke, Schmitz, Clausen, Hosenfeld, Köller 1997). Die Abstände zwischen den Lernenden der verschiedenen Schularten scheinen sich über die Schulzeit hinweg jedoch zu vergrößern. Dies konnte besonders für die Kompetenzen in Deutsch, in geringerem Ausmaß auch für mathematische und naturwissenschaftliche Fähigkeiten mehrfach belegt werden (bspw. Baumert et al. 2006b; Guill et al. 2010; Köller 1996; Scharenberg 2012). Besonders große Unterschiede in der Leistungsentwicklung zeigen sich zwischen Lernenden an Hauptschulen und Gymnasien, wobei die Leistungsentwicklung an Gymnasien am stärksten ist, gefolgt von Realschulen. Gesamt- und Hauptschulen weisen die schwächste Leistungsentwicklung auf (Baumert et al. 2006b). Dieser sogenannte Schereneffekt findet sich auch, wenn mögliche Effekte durch die soziale Herkunft oder die Vorleistung berücksichtigt werden (Becker et al. 2012; Guill et al. 2017; Pfof et al. 2010). Trotz der zahlreichen Studien, welche den Schereneffekt nachweisen, ist nicht davon auszugehen, dass der Schereneffekt kontinuierlich und über verschiedene Inhalte hinweg gegeben ist. Im Rahmen der KESS-Studie wurden über einen Zeitraum von zehn Jahren Lernende an Hamburger Sekundarschulen in verschiedenen Kompetenzdomänen getestet (Vieluf et al. 2011). Auch in dieser Studie konnte mehrfach der Schereneffekt repliziert werden, bspw. von der 4. zur 6. Klasse in Lesen und Mathematik sowie von der 7. zur 8. Klasse in Mathematik, Englisch und den Naturwissenschaften (Guill et al. 2010; Ivanov und Nikolova 2010; Nikolova und Ivanov 2010; Wendt et al. 2010). Für Orthographie und Lesen von der 7. zur 8. Klasse haben die Gymnasien hingegen den geringsten Leistungszuwachs (May 2010; Wendt et al. 2010). Besonders von der 8. zur 10. Klasse ist der oben beschriebene Schereneffekt in Lesen, Mathematik, Englisch und Naturwissenschaften nicht nachweisbar (Ivanov 2011a, 2011b; Nikolova 2011a, 2011b). Eine Erklärung hierfür könnten die teils hohen Ausgangsniveaus, vor allem an Gymnasien, sein. Ein

weiterer Grund könnten unterschiedliche Förderschwerpunkte sein. So verzeichnen zwar die Hauptschulen mit dem geringsten Ausgangsniveau in der 8. Klasse die größten Lernzuwächse in Lesen, Mathematik und Englisch. In den Naturwissenschaften ist der Leistungszuwachs jedoch am höchsten an Gesamtschulen.

Trotz des teilweise nachlassenden oder ausbleibenden Schereneffektes in höheren Klassenstufen ist über die gesamte Schulzeit in der Sekundarstufe I von einem Schereneffekt auszugehen. So können bspw. Baumert et al. (2006b) Schereneffekte in Mathematik und Englisch von der siebten bis zur zehnten Klasse nachweisen. Auch in der KESS-Studie zeigt sich der oben beschriebene Schereneffekt, wenn die Leistungsunterschiede zwischen der 4. und der 10. Klasse analysiert werden. Damit erfüllen die leistungsdifferenzierten Bildungsgänge nicht die Erwartungen von Befürwortern eines gegliederten Bildungssystems, welche dem gegliederten Schulsystem die optimale Förderung aller Lernenden in leistungshomogenen Gruppen zusprechen. Aber auch integrierte Gesamtschulen oder Schulen mit mehreren Bildungsgängen können nicht mit einer positiveren Leistungsentwicklung überzeugen.

Neben den Leistungen der Lernenden sind auch Unterschiede in den psychosozialen Merkmalen zwischen den Lernenden der verschiedenen Schularten von Interesse. Empirisch untersucht worden sind Merkmale wie Selbstwert, Leistungsangst, Schulunlust, Interesse, Motivation oder die Selbstwirksamkeit. Oftmals werden diese Merkmale neben der Leistung als Ergebnis von Unterricht untersucht. Darüber hinaus sind diese Merkmale für die Beschreibung des Unterrichts selbst von Bedeutung, da die Lernenden mit ihren psychosozialen Eigenschaften das Lernumfeld entscheidend prägen (siehe Abschnitt 3.4.3). Schulleiter und Schulleiterinnen an Hauptschulen, Schulen mit mehreren Bildungsgängen und integrierten Gesamtschulen berichten beispielsweise häufiger Beeinträchtigungen im Lernumfeld durch problematisches Schülerverhalten, mangelnde Arbeitshaltung und Störungen des Unterrichts als Schulleiter an Realschulen und insbesondere an Gymnasien (Hertel et al. 2010; Senkbeil et al. 2004). Zudem wird die Schüler-Lehrer-Beziehung aus Sicht der Lernenden an Schulen mit mehreren Bildungsgängen am ungünstigsten beurteilt, aber auch am Gymnasium scheint die Schüler-Lehrer-Beziehung relativ ungünstig zu sein (Hertel et al. 2010). Letzteres könnte mit der geringeren konstruktiven Unterstützung in Gymnasien einhergehen (siehe Abschnitt 3.5.3). Inwieweit diese Ergebnisse jedoch tatsächlich mit den affektiven Merkmalen der Lernenden in

Bezug stehen, ist auf Basis der bisherigen Forschungsrundlage schwer zu beantworten. Bezogen auf die Ausprägungen der affektiven Merkmale stellt Kunter (2005) in einer Varianzanalyse zwar fest, dass die Verteilungen der Merkmale Selbstkonzept, Interesse, Leistungsangst und Selbstwirksamkeit zu einem bedeutenden Teil auf die Schularten zurückzuführen sind (Kunter 2005, S. 174). Allerdings fallen die Unterschiede diesbezüglich zwischen den Schularten in den meisten Untersuchungen geringer aus als bei den Leistungen (bspw. Baumert et al. 2006b; Helmke, Schrader et al. 2008; Kunter 2005). Zudem ist die Befundlage nicht eindeutig. In manchen Studien zeigen die Lernenden an Gymnasien die höchsten Ausprägungen auf diesen Merkmalen, Hauptschulen die geringsten. Die anderen Schularten unterscheiden sich meist kaum voneinander (Baumert, Lehmann, Lehrke, Schmitz, Clausen, Hosenfeld, Köller 1997; Max-Planck-Institut für Bildungsforschung 1996). In anderen Studien werden entgegen der Erwartung die höchsten Werte bei Merkmalen wie Selbstkonzept, Interesse und Motivation bei den Lernenden an Hauptschulen gemessen, bei Lernenden an Gymnasien hingegen die niedrigsten Werte (Pekrun und Zirngibl 2004; Wagner et al. 2008).

Neben den Momentaufnahmen in der Einschätzung von psychosozialen Merkmalen interessiert besonders, wie sich die Lernenden im Laufe der Schulzeit entwickeln. Hier zeigen sich interessante Unterschiede zwischen den Schularten. Bei Lernenden an Gesamtschulen lässt sich bspw. ein Anstieg des Egoismus- bei simultanem Absinken des Altruismus-Motivs feststellen. Lernende an Hauptschulen entwickeln hingegen stärkere Konformitätsmotive (Max-Planck-Institut für Bildungsforschung 1996). Letzteres könnte eine Folge des enger geführten Unterrichts sein, den Lernende an Hauptschule erfahren, verglichen mit den anderen Schularten (siehe Abschnitt 3.5.3). Auch das Selbstkonzept entwickelt sich unterschiedlich zwischen den Schularten. Den stärksten Zuzugewinn in mathematisch-naturwissenschaftlichen Fächern verzeichnen Lernende an Hauptschulen, den geringsten Lernende an Gesamtschulen (Baumert, Lehmann, Lehrke, Schmitz, Clausen, Hosenfeld, Köller 1997; Becker und Neumann 2016; Helmke, Schrader et al. 2008). Die große Steigerung bei den Hauptschülern kann mit dem Big-Fish-Little-Pond-Effekt erklärt werden (siehe Abschnitt 3.4.4). Danach würde die – in der Regel – fähigkeitsangemessene Platzierung im traditionellen Bildungssystem positiv für die Entwicklung des Selbstkonzeptes sein. Eine gegensätzliche Entwicklung des akademischen Selbstkonzeptes findet sich hingegen für sprachliche Fächer im Rahmen der DESI-Studie. Hier verringert sich das akademische Selbstkonzept in Deutsch

und Englisch von Beginn zum Ende der neunten Klasse an Hauptschulen. An Gymnasien ist hingegen ein leichter Zugewinn zu verzeichnen. Eine generell negative Entwicklung scheint es beim Interesse der Lernenden zu geben. Zumindest für das Fach Deutsch finden Wagner et al. (2008) im Rahmen der DESI-Studie, dass das Lerninteresse sowie das Leseinteresse für das Fach Deutsch bei den Lernenden aller Schularten im Laufe des neunten Schuljahres abnimmt. Trotzdem zeigen sich hier Unterschiede zwischen den Schularten. Am geringsten nimmt das Interesse im Mittel an Gymnasien ab, am stärksten an den Hauptschulen. Realschulen und integrierte Gesamtschulen liegen im Mittelfeld (Wagner et al. 2008).

Insgesamt ist die Befundlage, die psychosoziale Merkmale zwischen Schularten beschreibt, jedoch bislang gering. Die hier beschriebenen Ergebnisse beziehen sich auf verschiedene Fächer, Klassenstufen und Entwicklungszeiträume. Es ist allerdings anzunehmen, dass die Entwicklung dieser Merkmale zum einen fachspezifisch ist, zum anderen auch von den erhobenen Zeiträumen abhängt (siehe Abschnitt 3.4.4). Dies könnte auch die teils variierende Befundlage zu den Ausprägungen und Entwicklungen der psychosozialen Merkmale erklären. Weitere Forschung speziell mit dem Fokus auf Unterschiede zwischen den Schularten wäre hier wünschenswert, um die überwiegend theoretisch angenommenen Unterschiede in den affektiven Merkmalen der Lernenden zwischen den Schularten empirisch besser abzusichern.

Die Lernenden prägen mit ihren Eigenschaften und Verhaltensweisen das Lernumfeld in der Schule und besonders im Klassenzimmer (siehe Abschnitt 3.4.3). Bestimmte Merkmalskonstellationen scheinen, wie der Forschungsüberblick zeigt, spezifische für die jeweiligen Schularten zu sein, beispielsweise die leistungsbezogene Zusammensetzung und die Zusammensetzung hinsichtlich sozio-ökonomischer Merkmale. Auch bei den affektiven Schülermerkmalen werden systematische Unterschiede zwischen den Schularten angenommen und teilweise empirisch nachgewiesen. Baumert und Kollegen (2001, 2006) folgern, dass sich hieraus spezifische Lern- und Entwicklungsmilieus ergeben, wodurch sie die Entwicklungsmöglichkeiten der Lernenden zwischen den Schularten unterscheiden.

„Wenn wir von differenziellen Lern- und Entwicklungsmilieus sprechen ist damit gemeint, dass junge Menschen unabhängig von und zusätzlich zu ihren unterschiedlichen persönlichen, intellektuellen, kulturellen, sozialen und

ökonomischen Ressourcen je nach besuchter Schulform differenzielle Entwicklungschancen erhalten, die schulmilieubedingt sind und sowohl durch den Verteilungsprozess als auch durch die institutionellen Arbeits- und Lernbedingungen und die schulformspezifischen pädagogisch-didaktischen Traditionen erzeugt werden“ (Baumert et al. 2006b, S. 98–99).

Theoretisch werden die Lern- und Entwicklungsmilieus bei Baumert und seinen Kollegen nicht tiefer fundiert. Am ehesten passt diese Verwendung des Milieubegriffes zu den Milieubeschreibungen von Taine (1828–1893) und Comte (1798–1857), welche Milieus als spezifische Umwelt- und Lebensverhältnisse von Personen bezeichnen. Die Milieufaktoren gelten dort als objektiv erfassbar und beeinflussen das Denken und Handeln in vorhersehbarer Weise (Gukenbiehl 2001). Beschrieben werden die Lern- und Entwicklungsmilieus über Faktoren der Institution, wie den Lehrplan, Lern- und Unterrichtskulturen und die Lehrerbildung sowie Faktoren der Schülergruppe wie die soziale, leistungsbezogene oder psychosoziale Zusammensetzung der Gruppe (Baumert und Köller 1998; Baumert et al. 2006b). Inhaltlich schließen diese Milieufaktoren damit an die eingangs beschriebenen Einflussgrößen aus dem erweiterten didaktischen Dreieck an (Abbildung 4): • Bildungsziele und -inhalte, • Lehrerbildung sowie schulartspezifischen Leitbilder • Selektion und Segregation. Analysiert wurden in den Studien zu den Lern- und Entwicklungsmilieus allerdings überwiegend Einflüsse der Schülerkomposition auf die Entwicklung der Lernenden in Abgrenzung zum Merkmal der Institution (operationalisiert über die Schulart). Es kann gezeigt werden, dass zusätzlich zur Schülergruppe auch die Institution mit der Leistungsentwicklung zusammenhängt (Baumert et al. 2006a; Baumert et al. 2009). Über die Schulart und die Komposition hinaus werden keine institutionellen Merkmale in die Analysen der differenziellen Lern- und Entwicklungsmilieus aufgenommen. Studien zu Kompositionseffekten, Lehrerüberzeugungen und Unterrichtsprozessen interpretieren die Ergebnisse jedoch häufig in Bezug auf schulartspezifische Lern- und Entwicklungsmilieus.

3.5.2 Lehrende und ihre Überzeugungen zum Lehren und Lernen

Bei der Beschreibung der Ausbildung Lehrender (Abschnitt 3.2) wurden Unterschiede in den Ausbildungsinhalten, den schulischen Zielen der verschiedenen Schularten und den professionellen Kompetenzen der Lehrenden herausgearbeitet. Dabei wurde argumentiert, dass Lehrende hierdurch schulartspezifische

Überzeugungen und Differenzen in den Unterrichtskulturen herausbilden. In diesem Abschnitt wird anhand der Forschung zu Überzeugungen von Lehrenden geprüft, inwieweit sich bislang solche systematischen Unterschiede zwischen den Schularten finden lassen. Insgesamt finden sich vergleichsweise wenige Studien, welche die Überzeugungen von Lehrenden differenziert nach den Schularten analysieren. Die meisten Studien untersuchen Zusammenhänge von Überzeugungen mit dem unterrichtlichen Handeln der Lehrenden (siehe Abschnitt 3.2). Wenige Ausnahmen, die auch die Schulart berücksichtigen, finden sich in PISA, DESI, TIMSS, BIJU und einer qualitativen Studie (Bohl et al. 2013).

Im Rahmen der DESI-Studie wurden Lehrende zu ihren Zielen im Deutsch- und Englischunterricht befragt. Hier zeigen sich deutliche Unterschiede in der Schwerpunktlegung zwischen den Schularten. Lehrende an Hauptschulen legen im Fach Deutsch den Schwerpunkt auf das freie „Sprechen“ und betonen die korrekte Aussprache von Wörtern, sowie die korrekte Wortwahl und Grammatik. Dieser Zielbereich hat auch an integrierten Gesamtschulen große Bedeutung, gemeinsam mit dem Ziel, Arbeitsmethoden wie Zeitmanagement, Lern-techniken und Informationsbeschaffung zu vermitteln. An Realschulen wird hingegen der größte Wert auf den schriftlichen Bereich gelegt bezogen auf Rechtschreibung und Grammatik. Die höchste Bedeutung für Gymnasiallehrkräfte hat hingegen die Vermittlung von Strategien der Texterarbeitung (Klieme, Eichler et al. 2008). Im Englischunterricht zeigen sich ebenfalls Unterschiede in den Unterrichtszielen. Hier legen jedoch die Lehrenden an Gymnasien den größten Wert auf die Kommunikation. In Hauptschulen spielt die Kommunikation in Englisch eine vergleichsweise geringe Rolle. Integrierte Gesamtschulen und Realschulen betonen im Vergleich zu den anderen Schularten das Übersetzen als besonders wichtig (Helmke, Helmke et al. 2008). Im Rahmen der DESI-Studie wurde ebenfalls die Leistungserwartung der Lehrenden vergleichend zwischen den Schularten untersucht. Hier unterscheiden sich die Fachkollegien der Schularten erwartungsgemäß. Insgesamt liegen die Leistungserwartungen im Gymnasium, und hier vor allem im Fach Englisch, deutlich höher als an den anderen Schularten. Nach dem gleichen Muster unterscheiden sich die Fachkollegien in der Wertschätzung der einzelnen sprachlichen Kompetenzen (Helmke, Helmke et al. 2008; Klieme, Eichler et al. 2008).

Die TIMS-Studie erfasst die Überzeugungen von Lehrenden zu Bedingungen für gute Schulleistungen in Mathematik und Naturwissenschaften. Bei den

Lehrenden besteht über die Schularten hinweg relative Einigkeit, dass Memorieren, konvergentes und divergentes Denken sowie die Fähigkeit, Inhalte kontextübergreifend anzuwenden, ähnlich wichtig seien. Es ergeben sich jedoch Unterschiede, wenn man die Antworten nach Schularten unterteilt. Lehrende an Gymnasien halten kreatives Denken für die wichtigste Fähigkeit, Lehrende an Hauptschulen betonen häufiger die praktische Anwendung (Baumert, Lehmann, Lehrke, Schmitz, Clausen, Hosenfeld, Köller et al. 1997).

Ähnliche Unterschiede zwischen den Überzeugungen von Lehrenden verschiedener Schularten liefert die Befragung von Lehrenden zu den Zielen im Mathematikunterricht und den präskriptiven Vorstellungen zum Mathematiklernen in PISA 2003. Betrachtet man die einzelnen Zieldimensionen im Vergleich zwischen den Schularten, sind keine Unterschiede für die motivationalen und metakognitiven Zieldimensionen („Wecken von Interesse“ und „Selbstregulationsfähigkeit“) zu erkennen (Kunter 2005). Im Hinblick auf die kognitiven Unterrichtsziele zeichnen sich allerdings charakteristische Schulformprofile ab. Lehrende an Schulen mit mehreren Bildungsgängen und Hauptschulen betonen für ihren Unterricht insbesondere die Beherrschung von Routinen und Standardaufgaben – Zielsetzungen, die an integrierten Gesamtschulen und Gymnasien geringer geschätzt werden. Gleichzeitig scheint an Hauptschulen und integrierten Gesamtschulen das Ziel der Vermittlung von Mathematisierungsfähigkeit zurückzutreten. Dieses Ziel stellen Lehrende an Gymnasien in den Vordergrund. Die Anwendung in Mathematik spielt wiederum an Hauptschulen eine deutlich größere Rolle als an Gymnasien (Baumert et al. 2004; Kunter et al. 2005).

Erfasst wurden in PISA 2003 zudem die impliziten Lerntheorien, d.h. die Vorstellungen von Lehrenden, wie Kinder und Jugendliche lernen. Hier werden fünf Dimensionen unterschieden: Selbständiges und diskursives Lernen, Vertrauen auf mathematische Selbständigkeit, Lernen durch Beispiele und Vormachen, Lernen durch Einschleifen sowie Lernen durch Reduktion auf Eindeutigkeit. Ein differenzierter Blick nach Schularten ergibt ein ähnliches Bild wie bei den Zielen im Mathematikunterricht. Vorstellungen vom selbständigen und diskursiven Lernen und ein höheres Vertrauen auf mathematische Selbständigkeit findet man vor allem bei Gymnasiallehrkräften, die einem rezeptiven Lernen an Beispielen und einem repetitiven Einschleifen gegenüber skeptisch sind. Dies sind Vorstellungen zum Mathematiklernen, die vor allem an Hauptschulen

und Schulen mit mehreren Bildungsgängen anzutreffen sind (Baumert et al. 2004; Kunter et al. 2005).

In einer qualitativen Studie (Bohl et al. 2013) wurden Lehrende in Interviews zu ihren Vorstellungen über das Lehren und Lernen befragt. In den Interviews nennen Lehrende an Gymnasien gegenüber Lehrenden anderer Schulformen häufiger Prinzipien zur kognitiven Aktivierung, insbesondere die Selbständigkeit bei der Bearbeitung von Aufgaben und Freiraum bei der Erarbeitung von Wissen. Lehrende an Hauptschulen betonen hingegen häufiger die Gestaltung von Aufgaben sowie die Lebensnähe von Themen. Daraus ziehen Bohl et al. (2013) den Schluss, dass Lehrende an Hauptschulen Aspekten des kognitiven Kompetenzerwerbs weniger Bedeutung zusprechen als Lehrende an Gymnasien (Bohl et al. 2013, S. 103).

Die häufig in der angloamerikanischen Forschung untersuchte These, dass Lehrende ihr Handeln dem Leistungsniveau und der sozialen Zusammensetzung der Klasse anpassen (siehe Abschnitt 3.4.5), konnte basierend auf Daten der BIJU-Studie für das deutsche Bildungssystem innerhalb der Schulformen nicht belegt werden. Gruehn (2000) findet auf der Makroebene eine Anpassung des Unterrichtsstils entsprechend den Schulformen. In nicht gymnasialen Schulen wird mehr repetitiv geübt, in gymnasialen Klassen spielen konstruktivistische Ansätze eine größere Rolle. Innerhalb der Schularten erweist sich die Verwendung lernförderlicher bzw. lernmindernder Unterrichtsstrategien als unabhängig vom kognitiven Ausgangsniveau (Gruehn 2000, S. 210). Hieraus lässt sich die vorsichtige These ableiten, dass schulartspezifische Überzeugungen eine bedeutendere Rolle für die Gestaltung des Unterrichts spielen als Überzeugungen bezüglich der tatsächlich unterrichteten Schülergruppen.

Zusammenfassend lassen sich auf Basis der hier beschriebenen Studien, die sich mit Unterschieden in den Überzeugungen der Lehrenden befassen, Unterschiede zwischen den Schularten hinsichtlich der Überzeugungen, wie Unterricht in Bezug auf das kognitive Anregungspotential gestaltet wird, erwarten; besonders zwischen Gymnasien und Hauptschulen. An Hauptschulen werden das Üben, Wiederholen, Routinieren und der Alltagsbezug betont. In Gymnasien stehen kreatives Denken, Selbständigkeit und Modellieren stärker im Fokus. Hinsichtlich motivationaler und metakognitiver Dimensionen (wie dem Ziel, Interesse zu vermitteln,) scheinen sich die Überzeugungen der Lehrenden zwischen den Schularten jedoch weniger zu unterscheiden. Oftmals werden in der

Literatur die Ergebnisse allerdings nur für die Schularten Hauptschule und Gymnasium interpretiert oder das Gymnasium den anderen Schularten gegenübergestellt. Wünschenswert wären hier getrennte Beschreibungen der Schularten; auch um klarer herauszufiltern, welche Merkmale sich nicht zwischen einzelnen Schularten unterscheiden. Von besonderem Interesse ist auch, wie sich die schulartspezifischen Überzeugungen in entsprechendem Lehrerhandeln niederschlagen. Dies wird in den überwiegend deskriptiven Studien kaum thematisiert.

3.5.3 Gegenstand: Unterrichtszeit und Unterrichtsinhalte

Der Gegenstand des Unterrichts, hier spezifiziert als die Unterrichtszeit und Unterrichtsinhalte in einzelnen Unterrichtsfächern (siehe Abschnitt 3.3), ist nur selten Thema von Schulartvergleichen. Besonders die Unterrichtszeit scheint vergleichend zwischen den Schularten wenig zu interessieren. Diese Daten können, zumindest für das intendierte Curriculum, relativ einfach den Stunden- tafeln in Schulordnungen der Länder entnommen werden (bspw. Bayerisches Staatsministerium für Unterricht und Kultus 2007; Hessisches Kultusministerium 2011d; Sächsisches Staatsministerium für Kultus; Thüringer Ministerium für Bildung 2018). Für Kernfächer wie Mathematik sind die vorgeschriebenen Unterrichtszeiten über die Länder und Schularten hinweg in den meisten Fällen für die 9. Klasse auf 4 Unterrichtsstunden (3 Zeitstunden) festgelegt. In wenigen Ausnahmen sind 5 Unterrichtsstunden (3,75 Zeitstunden) verpflichtend, bspw. für Haupt- und Realschulen in Bayern, oder 3 Unterrichtsstunden (2,25 Zeitstunden), bspw. in Regelschulen in Thüringen. Anhand von Schülerangaben in PISA 2003 finden Drechsel und Senkbeil (2004) ähnliche, jedoch nicht identische Angaben. Lernende an Hauptschulen geben im Mittel 3,7 Zeitstunden an. Lernende an integrierten Gesamtschulen und Realschulen 3,1 bzw. 3,0 Zeitstunden. Gymnasiasten geben die geringste Anzahl an Stunden an (2,7 Zeitstunden). An diesem Beispiel wird die Diskrepanz zwischen intendiertem und realisiertem Curriculum deutlich. Wobei für die Stundentafel in Mathematik besonders an Hauptschulen von den Vorgaben abgewichen wird, was die vergleichsweise hohe Varianz ($SD=2,1$) verdeutlicht. Gründe für mehr Mathematikunterricht könnte bspw. zusätzlicher Förderunterricht sein. Gleichzeitig scheinen andere Lernende an Hauptschulen deutlich weniger als 3 Zeitstunden pro Woche in Mathematik unterrichtet zu werden. Erstaunlich ist zudem, dass auch Lernende am Gymnasium – relativ übereinstimmend ($SD=0,7$) – weniger

Mathematikunterricht besuchen als vorgeschrieben. Hier könnten Schulartunterschiede in den Ursachen für die Abweichungen in den Unterrichtsstunden vorliegen. Aus der aktuellen Datenlage lassen sich diese jedoch nicht identifizieren. Deshalb bleibt vorerst die Schlussfolgerung, dass für den Mathematikunterricht die Vorgaben in den Stundentafeln nahezu identisch sind. In den tatsächlich unterrichteten Stunden weichen die Schularten im Mittel, aber vor allem in der Varianz sichtlich voneinander ab.

Unterrichtsinhalte werden in den verschiedenen Studien sehr unterschiedlich erhoben. Dies kann neben den differierenden Studienschwerpunkten durchaus auch durch die verschiedenen untersuchten Unterrichtsfächer bedingt sein, in denen die verwendeten Materialien, die Art der Bearbeitung und die Aufgabengestaltung durchaus von unterschiedlicher Relevanz sind. In der DESI-Studie wurde das im Deutschunterricht eingesetzte Aufgabenmaterial im Lehrerfragebogen erfasst. Hier zeigen sich Unterschiede in den behandelten Textsorten zwischen den Schularten. Während an Hauptschulen und integrierten Gesamtschulen vielfach Alltagsmaterialien wie Zeitungstexte, Tabellen und Schaubilder, aber auch berufsbezogene Schreiben ihren Platz im Deutschunterricht haben, nutzen die Gymnasien relativ häufig auch literarische Texte. Die breiteste Verwendung unterschiedlicher Materialien findet sich an integrierten Gesamtschulen (DESI-Konsortium 2008; Klieme, Jude et al. 2008). Für das Fach Geschichte hat Martens (2010) den Umgang mit Geschichtsdarstellungen untersucht. Dazu wurde die Diskussion von zwei kontroversen Darstellungen eines historischen Ereignisses mit der dokumentarischen Methode (vgl. Bohnsack 2007) analysiert. Hier zeigte sich, dass in den Diskussionen an Haupt- und Realschulen die Differenzen zwischen den Texten marginalisiert wurden und die Diskussion stark an der eigenen Lebenswelt und dem eigenen Geschichtsbild orientiert war. Die Diskussionen an Gymnasien waren hingegen stärker reflektiert im Umgang mit den Informationen und es wurden verschiedene Ursachen der kontroversen Darstellungen diskutiert, bspw. die verwendeten Quellen oder das Erkenntnisinteresse der Autoren (Martens 2010). Hieraus lässt sich vorsichtig folgern, dass der Geschichtsunterricht an Gymnasien eher reflektierend und analytisch ist, entsprechend dem wissenschaftspropädeutischen Bildungsziel dieser Schulart. An Haupt- und Realschulen scheint hingegen die Vermittlung von Geschichte als Faktenwissen vordergründig. Auch für das Fach Mathematik finden sich Unterschiede zwischen den Schularten, welche mit den Bildungszielen der Schularten und den Überzeugungen der Lehrenden einhergehen. In der COACTIV-Studie haben Neubrand et al. (2011) die

von den Lehrenden eingesetzten Mathematikaufgaben analysiert und auf Unterschiede in den curricularen Niveaus vergleichend zwischen Gymnasien und den anderen Schularten untersucht. Während in Gymnasien eher innermathematische Bezüge hergestellt werden, findet sich an den anderen Schularten ein deutlich höherer Anteil an Aufgaben mit Anwendungsbezug zum Alltag (Neubrand et al. 2011, S. 129). Die Befunde zu den eingesetzten Aufgaben stimmen somit gut mit den Überzeugungen der Lehrenden zu den Zielen im Mathematikunterricht überein (siehe Abschnitt 3.5.2).

Zusammenfassend unterscheiden sich die Aufgabeninhalte in verschiedenen Fächern zwischen den Schularten in den verwendeten Materialien, der Art der Bearbeitung und der Aufgabengestaltung. Die unterschiedlichen Schwerpunktlegungen lassen sich gut über die schulartspezifischen Bildungsziele und die Überzeugungen der Lehrenden herleiten. Empirische Analysen, welche diese bislang eher hypothetischen Zusammenhänge prüfen, wären hier wünschenswert, um die Bedeutung der schulartspezifischen Bildungsziele und Lehrerüberzeugungen für die Wahl und Gestaltung der Unterrichtsinhalte zu ergründen.

3.5.4 Unterrichtsprozesse

„Die hier vorgelegten Ergebnisse eines Forschungsprojektes zu Fragen der Unterrichtsmethoden sollen zur Erhellung eines zentralen unterrichtstheoretischen und -praktischen Problembereichs beitragen, der nach unserer Ansicht bislang nicht den Stellenwert im erziehungswissenschaftlichen Forschungsbetrieb einnahm, den er aufgrund seiner Bedeutung für Lehrende und Lernende haben müsste“ (Hage et al. 1985, S. 4). Dieser Kritik, die Hage et al. (1985) am Forschungsfeld zum Unterricht äußern, kann aus heutiger Sicht nicht mehr stattgegeben werden, ganz im Gegenteil. Nicht zuletzt mit der „Hattie Studie“ (Hattie 2009) erfreut sich die Erforschung des Unterrichts und des Lehrerhandelns großer Beliebtheit. Für den Unterricht der Lernenden findet sich deshalb eine große Bandbreite an untersuchten Konzepten aus verschiedenen Forschungstraditionen, welche in dieser Arbeit nicht erschöpfend aufgearbeitet werden können. Der Forschungsüberblick fokussiert daher überwiegend die Ergebnisse aus der empirischen Lehr-Lernforschung der groß angelegten Leistungs- und Vergleichsstudien. Die Merkmale sind in den verschiedenen Studien nicht immer klar nach Struktur- und Prozessmerkmalen sowie nach Inhalten

(bspw. von Unterrichtsaufgaben) trennbar. Bestes Beispiel ist die kognitive Aktivierung, welche als Merkmal der Unterrichtsqualität zu den Prozessmerkmalen gezählt werden kann (bspw. Kunter und Trautwein 2013). Dieses Merkmal kann aber auch im Sinne von Methoden zur kognitiven Aktivierung betrachtet werden oder hinsichtlich des Aufgabenmaterials zur kognitiven Aktivierung (bspw. Bohl et al. 2013; Neubrand et al. 2011). Der Forschungsüberblick betrachtet die diversen Merkmale zur Beschreibung des Unterrichts daher gemeinsam. Studien, welche das Aufgabenmaterial hinsichtlich des Potentials zur kognitiven Aktivierung beschreiben, werden hier mit aufgeführt. Analysen der Unterrichtsinhalte und Unterrichtsmaterialien finden sich in Abschnitt 3.5.3.

Die historisch anmutende Untersuchung zum Schulunterricht von Hage et al. (1985) verdeutlicht die mangelnde Trennschärfe zwischen den Konzepten der Struktur- und Prozessebene eindrucksvoll, ebenso die sich wandelnde Bedeutungszuschreibung (bspw.: „Bei Klassenkooperation findet das Lernen zwar noch im Gesamtverband statt, der Lehrer befindet sich jedoch als Leiter im Hintergrund“; Hage et al. 1985, S. 38). Die fast 40 Merkmale umfassende Analyse zum Unterricht der Schularten Hauptschule, Gymnasium und Gesamtschule wird deshalb hier nur zusammenfassend behandelt. Über Unterrichtsbeobachtung, Interviews mit Lehrenden und standardisierte Fragebögen für Lehrende und Lernende wurde der Unterricht in 7. und 8. Klassen umfassend hinsichtlich des Unterrichtsstils, des Klimas und der Interaktionen beschrieben. Allerdings sei darauf hingewiesen, dass die Stichprobe mit zehn untersuchten Schulen relativ klein ist. Als relevant festzuhalten bleibt, dass sich die Schularten in den meisten untersuchten Merkmalen bedeutend unterscheiden, besonders zwischen Gymnasium und Hauptschule: „Die Schulformunterschiede waren dabei deutlicher [Anm. d. A.: als Fachunterschiede]. Zunächst fällt auf, dass die Unterschiede zwischen Gymnasium und Hauptschule besonders häufig am größten waren, während Gesamtschulen eine Mittelposition einnahmen. Das Gymnasium bezeichnen wir deshalb pointierend als Schule des ‚gelenkten Sachdialogs‘, die Hauptschule als Schule der ‚gelenkten Beschäftigung‘, die Gesamtschule als Schule des ‚gelenkten Diskurses‘“ (Hage et al. 1985, S. 148). Die Autoren identifizieren schulartspezifische Unterrichtskulturen, welche auch

mit dem heutigen Bild der Schularten vereinbar sind, wie der untenstehende Forschungsüberblick zu aktuelleren Studien zeigt.¹²

Relativ eindeutige Schulartunterschiede finden sich in der kognitiven Aktivierung, der Klassenführung und der konstruktiven Unterstützung. Dies ist zumindest dann der Fall, wenn bei der Analyse des Merkmals kognitiver Aktivierung die Erhebungsperspektive berücksichtigt wird (Clausen 2002). Bezüglich der kognitiven Aktivierung und in geringem Ausmaß auch bezüglich der konstruktiven Unterstützung berichten Lernende und Lehrende abweichende Einschätzungen. Aussagen von Lernenden zum Unterricht geben die Perspektive des erlebten oder wahrgenommenen Unterrichts wieder, die Aussagen von Lehrenden beziehen sich stärker auf die didaktische Perspektive (Clausen 2002; Kunter und Trautwein 2013).

Aus Sicht von Lehrenden bezüglich des Mathematikunterrichts findet ein kognitiv aktivierender Umgang mit mathematischen Inhalten eher in Gymnasien statt, während vor allem an Hauptschulen Unterrichtsformen mit geringem kognitivem Gehalt und ausgeprägter Unterstützung überwiegen. Die anderen Schulen befinden sich im mittleren Bereich. Aus der Schülerperspektive ist der Unterricht besonders an Hauptschulen kognitiv anstrengend, weniger jedoch an Gymnasien. (Baumert et al. 2004; Bayer et al. 2015; Gruehn 2000; Kunter und Voss 2011; Rakoczy et al. 2008). Die Lernenden an Hauptschulen fühlen sich vermutlich eher angestrengt, wohingegen Gymnasiasten bei dem gebotenen Anspruchsniveau gut arbeiten können. Teilweise könnte der Unterricht an Gymnasien – in Anbetracht der niedrigen Werte aus Sicht der Lernenden – im Anspruch noch erhöht werden. Auch im eingesetzten Aufgabenmaterial zeigen sich Unterschiede zwischen den Schularten. In einer Videoanalyse des Mathematikunterrichts in verschiedenen Schularten (Bohl et al. 2013) zeigt sich, dass die eingesetzten Aufgaben am Gymnasium komplexer sind als die Aufgaben in den untersuchten Real- und Hauptschulen. Ebenso konnte ein kognitiv-aktivierendes Handeln der Lehrpersonen in den Phasen der Aufgabeneinfüh-

12 Der Forschungsüberblick stellt die Befundlage aufgrund der Vielfalt und häufigen Überschneidung thematisch dar. Die Erkenntnisse aus den einbezogenen Studien, welche überwiegend auf Analysen der PISA-, TIMSS-, BIJU- und DESI-Daten beruhen, werden daher nicht nach Studien, sondern nach inhaltlichen Erkenntnissen zusammengefasst. Die Literaturgrundlage für die zusammenfassende Beschreibung wird daher teilweise am Ende eines thematischen Absatzes zitiert.

rung, der Begleitung der Aufgabenbearbeitung und der Besprechung der Aufgaben gefunden werden. Besonders die Unterschiede zwischen Gymnasium und Hauptschule sind sowohl bei der Aufgabenkomplexität als auch beim kognitiv-aktivierenden Handeln hervorstechend. In den beobachteten Unterrichtsstunden wurde zudem ein höherer Anteil an Besprechungsphasen im gymnasialen Unterricht gefunden (25,5 %) im Vergleich zu Realschulen (17,5 %) und Hauptschulen (9,9 %; Bohl et al. 2013, S. 103). Im Rahmen der COACTIV-Studie wurden die von den Lehrenden eingesetzten Mathematikaufgaben hinsichtlich der Unterschiede zwischen Gymnasien und den anderen Schularten in den curricularen und kognitiven Niveaus analysiert (Neubrand et al. 2011). Die Auszählung des untersuchten Materials ergab, dass argumentative Aufgaben an Gymnasien deutlich häufiger gestellt werden (1 von 10 Aufgaben) als an Schulen der anderen Schularten (1 von 50).

Konstruktive Unterstützung und Klassenführung werden von den Lernenden und Lehrenden für den Mathematikunterricht annähernd übereinstimmend beurteilt. Lernende an Hauptschulen berichten von einer deutlich höheren konstruktiven Unterstützung als Lernende an Gymnasien. Die anderen Schulformen befinden sich mit relativ ähnlichen Ausprägungen zwischen Hauptschule und Gymnasium. In der Klassenführung berichten Lehrende wie Lernende die niedrigsten Ausprägungen für das Merkmal in Hauptschulen, hier wird auch häufiger von Disziplinproblemen und Zeitverlust berichtet. An den anderen Schulformen wird die Klassenführung deutlich positiver eingeschätzt, besonders am Gymnasium. Am Gymnasium berichten die Lernenden entsprechend auch die geringsten Disziplinprobleme. Die anderen Schularten befinden sich wiederum mit relativ ähnlichen Ausprägungen zwischen Gymnasium und Hauptschule. Insgesamt wird die Angebotsstruktur des Mathematikunterrichts von Jugendlichen an Hauptschulen positiv beurteilt, während Gymnasiasten ihren Unterricht eher kritisch bewerten. Diese Differenzen sind vor allem durch Unterschiede in der individuellen Lernunterstützung durch die Lehrenden zu erklären, welche von Lernenden an Hauptschulen als besonders ausgeprägt, an Gymnasien eher als gering beschrieben werden. Die Hauptschulen zeichnen sich damit durch eine größere konstruktive Unterstützung aus, Gymnasien hingegen durch stärkere kognitive Aktivierung (Baumert et al. 2004; Baumert und Kunter 2006; Bayer et al. 2015; Gruehn 2000; Klieme und Rakoczy 2001; Klieme, Schümer et al. 2001; Tillmann und Meier 2001). Solche Profile entsprechen durchaus schulartspezifischer Didaktik und verdeutlichen, dass Schul-

arten in unserem Bildungssystem als differenzielle Entwicklungsmilieus wirken können (Klieme, Schümer et al. 2001).

Für den Deutschunterricht ergeben sich ähnliche Muster in der Unterrichtsbeschreibung wie oben für den Mathematikunterricht beschrieben. Allerdings liegen die Lernenden aus Hauptschulen und integrierten Gesamtschulen in ihrer Beschreibung des Unterrichts bezüglich der kognitiven Aktivierung und der konstruktiven Unterstützung näher beieinander. Lernende an Hauptschulen und an integrierten Gesamtschulen berichten übereinstimmend von zu hohem Tempodruck, aber auch von unterstützendem und gut strukturiertem Unterricht und von einer hohen Vielfalt an Lernsituationen. Lernende an Gymnasien berichten von einer geringen Vielfalt an Lernsituationen und wenig forderndem Unterrichtstempo (Hertel et al. 2010; Klieme, Jude et al. 2008). Ähnliche Unterrichtsbeschreibungen liegen auch für das Fach Englisch vor. Die Unterschiede sind jedoch deutlich geringer ausgeprägt. Ein hervorstechender Unterschied zwischen den Schularten ist die vergleichsweise hohe Bedeutung der Kommunikation an Gymnasien, welche besonders an Hauptschulen nur eine geringe Rolle spielt. Stärker als im Fach Deutsch unterscheidet sich im Fach Englisch der wahrgenommene Tempodruck zwischen den Schularten. Dieser wird an Hauptschulen als besonders hoch eingeschätzt. An Gymnasien wird das hohe Anspruchsniveau von den Lernenden betont (Helmke, Helmke et al. 2008).

Die Unterschiede zwischen den Schularten in der kognitiven Aktivierung, der konstruktiven Unterstützung sowie in der Klassenführung sind über mehrere Studien sowie über die Fächer hinweg relativ eindeutig. Trotzdem handelt es sich bei den hier beschriebenen Ergebnissen überwiegend um Ergebnisse aus Fragebogenstudien, die anfällig für Antworttendenzen sind, welche bspw. spezifisch für bestimmte Gruppen sein können. Besonders für die Klassenführung und die konstruktive Unterstützung mangelt es an alternativen Erhebungen wie Unterrichtsbeobachtungen, welche Vergleiche zwischen Schularten anstellen. Einen Hinweis für gruppenspezifische Antworttendenzen in Fragebogenerhebungen liefert die Studie von Klieme et al. (2010). Unter Kontrolle der sozialen und der ethnischen Komposition sind die Merkmale kognitive Aktivierung, konstruktive Unterstützung und Klassenführung aus Schülersicht an Schulen mit Hauptschulgang deutlich positiver beurteilt als an Realschulen oder Gymnasien. Aus diesem Grund zeigen sich hier für die Klassenführung abweichende Ergebnisse zu obengenannten Studien, welche nicht für Schülermerkmale kontrolliert haben. Hier besteht weiterer Forschungsbedarf, um die Verzerrung der

Daten durch gruppenspezifische Antworttendenzen einschätzen bzw. die Merkmale bereinigt um Antworttendenzen zwischen den Schularten untersuchen zu können.

Weniger systematische Varianz zwischen den Schularten zeigt sich hinsichtlich der eingesetzten Unterrichtsmethoden. In den verschiedenen Studien sind die Merkmale zum einen sehr unterschiedlich erhoben. Zum anderen sind die methodischen Zugänge in der heutigen Zeit anscheinend weniger von der Schulart abhängig. Für den Deutschunterricht wird im Rahmen von DESI beispielsweise aus den Daten zu verschiedenen Organisationsformen des Unterrichts gefolgert, dass diese in allen Schulformen Einzug erhalten haben (Klieme, Jude et al. 2008). Auch Kunter und Voss (2011) finden in der COACTIV-Studie zwar individuelle Lehrstile beim Einsatz verschiedener Lern- und Sozialformen im Mathematikunterricht. Allerdings lassen sich diese nur bedingt durch die Schulart erklären. Lehrende an Hauptschulen und Gesamtschulen setzen zwar häufiger individuelle Arbeitspläne ein, insgesamt sind solche Effekte aber sehr gering. Dies führt sie zu der Schlussfolgerung, dass ein übergreifendes Muster an Sichtstrukturen im Mathematikunterricht vorherrscht, das relativ wenig zwischen verschiedenen Schulformen variiert (Kunter und Voss 2011, S. 96). Für ausgewählte Merkmale finden Schiepe-Tiska et al. (2016) im naturwissenschaftlichen Unterricht Unterschiede im Vergleich zwischen Gymnasien mit den anderen Schularten. In Gymnasium wird der Unterricht öfter an die Bedürfnisse und das Wissen der Lernenden angepasst. Zudem wird im naturwissenschaftlichen Unterricht mehr diskutiert und die Lernenden zum Erklären angeregt. In den anderen Schularten wird stattdessen häufiger experimentiert und ein stärkerer Bezug zum Alltag hergestellt (Schiepe-Tiska et al. 2016). Deutlicher als bei der vergleichenden Betrachtung einzelner Methoden zeigen sich Unterschiede in der übergreifenden Orientierung des Unterrichts zwischen den Polen der Schülerorientierung und der Lehrerorientierung (siehe Abschnitt 3.1.1). Hier findet Kunter (2005), dass der lehrerorientierte Unterricht besonders häufig am Gymnasium, der schülerorientierte Unterricht dagegen mehrheitlich an Hauptschulen oder integrierten Gesamtschulen zu beobachten war. Die Realschule nimmt eine Mittelposition ein. Hier sind lehrergesteuertes und schülerorientiertes Vorgehen annähernd ausgeglichen.

Zusammenfassend finden sich Unterschiede zwischen den Schularten in den generischen Faktoren des Unterrichts und den übergreifenden Orientierungen. Der Einsatz der Methoden scheint sich jedoch stärker an den Themen und Fä-

chern zu orientieren, als dass die Schulart die bedeutendere Rolle für die Auswahl der Methoden spielt. Allerdings lässt sich in der Forschungsliteratur ein Mangel an Differenzierung in der Analyse und Interpretation der Ergebnisse feststellen, da oftmals Vergleiche nur zwischen Gymnasium und den anderen Schulen bzw. zwischen Gymnasium und Hauptschule angestellt oder interpretiert werden. Zudem basieren die meisten Ergebnisse auf Fragbogendaten, die anfällig für Antworttendenzen sind und zu verzerrten Ergebnissen führen können. In den zitierten Studien konnten keine Hinweise gefunden werden, dass potentielle Verzerrungen durch Antworttendenzen untersucht und berücksichtigt wurden. Auch weitere Vorannahmen über die Vergleichbarkeit der Schularten wurden nicht geprüft (siehe Abschnitt 3.5). Über die Studien hinweg sind zudem teils erhebliche Abweichungen in der Operationalisierung gleicher oder ähnlicher Merkmale sowie unterschiedliche Schwerpunktsetzungen in den verglichenen Inhalten evident. Dies schränkt die Aussagekraft gefundener Schulartunterschiede im Unterricht stark ein. Untersuchungen, welche mit breiterer Sicht auf den Unterricht Schulartvergleiche anstellen und auch Zusammenhänge zwischen den Unterrichtsprozessen und damit zusammenhängenden Merkmalen, wie den Überzeugungen der Lehrenden oder Merkmalen der Schülergruppe, vergleichend untersuchen, fehlen bislang. Um diesen Forschungsbedarf zu bedienen, untersucht die vorliegende Arbeit systematisch Unterschiede der Zusammensetzung der Schülergruppen, der Zielvorstellungen der Lehrenden, der Unterrichtsinhalte sowie der Unterrichtsprozesse basierend auf einer repräsentativen Stichprobe für die 9. Klassen in Deutschland. Zudem wird untersucht, inwieweit die Unterschiede zwischen den Schularten in den Unterrichtsprozessen über die Zielvorstellungen der Lehrenden, die Eigenschaften der Schülergruppe und die Inhalte des Unterrichts erklärt werden können. Abschließend werden die Unterschiede in den Unterrichtsprozessen hinsichtlich ihres Zusammenhangs mit der Leistungsentwicklung der Lernenden analysiert.

4 Herausforderungen des Vergleichs von Schularten

Vergleiche in den Erziehungswissenschaften haben eine längere Tradition als es in den aktuellen Diskussionen um Vergleiche von Bildungssystemen und Globalisierung den Anschein hat (bspw. Adick 2005; Knobloch 2014). Berichte über Bildungseinrichtungen in anderen Ländern und Gesellschaften finden sich schon in den Aufzeichnungen von Entdeckungsreisen und Eroberungsfeldzügen (Allemann-Ghionda 2004, S. 18). In Zeiten der Eroberungszüge von Cortez entstanden bspw. detaillierte Beschreibungen über die aztekische Gesellschaft und das Sprachenlernen der aztekischen Kinder (Todorov und Böhringer 2008). Diese frühen Zeugnisse über fremde Bildungssysteme waren in der Regel jedoch eher narrativ als systematisch und Vergleiche mit dem eigenen Bildungssystem eher implizit als explizit (Allemann-Ghionda 2004, S. 22). Als Gründungsvater systematischer Datenerfassung und des Vergleichs im Bildungsbereich gilt Marc-Antoine Jullien de Paris (1775–1848), der standardisierte Fragebögen vorschlug, um Bildungsinstitutionen der verschiedenen europäischen Länder zu vergleichen. Einen deutlichen Aufschwung erfuhr der Vergleich von Bildungssystemen aber erst in der zweiten Hälfte des 19. Jahrhunderts. Fortschritte in den quantitativen Methoden führten zu strukturierten Studien und zu einem stärkeren Fokus auf Interpretationen im Sinne kultureller Eigenheiten, in dem einzelne Merkmale kontrastierend gegenübergestellt wurden (Adick 2005; Allemann-Ghionda 2004). Der Vergleich von sozialistischen mit demokratischen Bildungssystemen (bspw. Anweiler 1969) oder kollektivistischen mit individualistischen Kulturen (im Überblick bei Weippert 1964) sind Beispiele hierfür.

In der Erziehungswissenschaft werden solche Vergleiche kontrovers diskutiert. Besonders der Kulturbegriff wird kritisiert, da dieser schwer eingrenzbar sei und häufig unreflektiert gebraucht werde. Zum einen existieren oder wirken Kulturen über nationale Grenzen hinaus (Adick 2005; Schubert 2014): Klassische Konzepte aus der interkulturellen Bildungsforschung sind beispielsweise motivationale Fragen bezüglich Schule und Lernen (bspw. Baumert, Lehmann, Lehrke, Schmitz, Clausen, Hosenfeld, Köller 1997; Reiss et al. 2016). Werden diese zwischen den Ländern verglichen, könnten hieraus Motivations- und Einstellungskulturen interpretiert werden (bspw. OECD 2013). Naheliegender ist hingegen, dass ähnliche Muster über Ländergrenzen hinweg existieren. Zum

anderen können innerhalb einer Nation verschiedene kulturelle Traditionen koexistieren (Allemann-Ghionda 2004; Schubert 2014), so auch verschiedene Muster von Motivationen und Einstellungen von Lernenden innerhalb eines Landes. Neben der schwierigen Definition kultureller Grenzen stehen auch die Analysen in der Kritik, da Statistiken Unterschiede zwischen Ländern und Kulturen künstlich herstellten oder verstärkten. Variationen innerhalb der untersuchten Gruppen würden vernachlässigt und Zusammenhänge stark vereinfacht. Der Vergleich selbst produziert somit stets Vereinfachungen, da verschiedenartige Strukturen, Situationen und Arrangements auf gemeinsame Nenner reduziert werden (Allemann-Ghionda 2004, S. 155).

Auch die hier vorliegende Arbeit bewegt sich im Spannungsfeld der vergleichenden Forschung. Zwar werden hier nicht Nationen oder Kulturen sondern Schularten gegenübergestellt. Trotzdem werden diese entsprechend der theoretischen Herleitung als unterschiedliche und abgrenzbare Beschulungsstrukturen betrachtet, welche hinsichtlich inhaltlicher, struktureller und prozeduraler Differenzen untersucht werden. Ein solcher Vergleich bringt ähnlich der internationalen und interkulturellen Forschung die Gefahr der Vereinfachung und Zuschreibung mit sich, bspw. wenn Zusammenfassungen wie bei Hage et al. (1985, S. 148) unreflektiert verinnerlicht werden: „Das Gymnasium bezeichnen wir [...] pointierend als Schule des ‚gelenkten Sachdialogs‘, die Hauptschule als Schule der ‚gelenkten Beschäftigung‘, die Gesamtschule als Schule des ‚gelenkten Diskurses‘“ (siehe hierzu auch Kapitel 3.5.4). Auch Baumerts „Lern- und Entwicklungsmilieus“, welche direkt mit den Schularten in Bezug gebracht werden (Baumert et al. 2006a, S. 98–99) können zur Pauschalisierung verleiten.

Den Gefahren der vergleichenden Forschung stehen Argumente gegenüber, welche die negative Sicht abschwächen und die Vorteile von vergleichender Bildungsforschung betonen. So wird abschwächend gegen die Problematik künstlicher bzw. starrer kultureller und nationaler Grenzen in der vergleichenden Forschung argumentiert, dass kulturelle Differenzen nicht völlig im „transnationalen Blickwinkel“ (siehe hierzu Adick 2005) aufgehen. Das heißt, Bildungssysteme entstehen und entwickeln sich im Kontext gesellschaftlicher Systeme (Schubert 2014), auch wenn die Grenzen teilweise fließend sind. Werden diese Grenzen – bspw. bei Konzepten wie Transkulturalität oder Transnationalität, die von einer wechselseitigen Durchdringung von Kulturen in Prozessen der Modernisierung oder Globalisierung ausgehen, – ignoriert statt proble-

matisiert, würde damit Universalität simuliert (Allemann-Ghionda 2004). In der aktuellen Bildungsforschung werden bspw. allzu oft Unterrichtskonzepte und Praktiken empfohlen, ohne dass Unterschiede in der Ausgangssituation diskutiert werden. Ein Klassiker ist hierbei die Studie „Invisible Learning“ von John Hattie (2009), welche Merkmale identifiziert, die den kognitiven Lernerfolg von Lernenden positiv bedingen sollen. Dass Lehrerhandeln jedoch nicht verallgemeinert den gleichen Effekt bei den Lernenden haben muss, lässt sich aus den Ergebnissen ländervergleichender Forschung folgern: „most teaching and learning strategies do not have a direct, robust and consistent relationship with student performance across countries“ (OECD 2010, S. 12). Auch bezogen auf die Schularten in Deutschland liefern Zusammenhangsanalysen von Unterrichtsmerkmalen und Schülerleistung unter Berücksichtigung der Schulform bei Gruehn (2000) Hinweise auf unterschiedliche Zusammenhänge zwischen Unterrichtsmerkmalen und der Schülerleistung. Sie folgert aus den Ergebnissen, dass repetitives Üben in der Hauptschule einen kompensatorischen Charakter hat und somit zu größerem Lernerfolg führt. An Gymnasien förderten hingegen konstruktivistische Unterrichtsansätze stärker die Leistungsentwicklung. Pauschale Empfehlungen von konstruktivistischen Unterrichtsansätzen würden nach diesen Ergebnissen Lernende an Hauptschulen benachteiligen, weshalb differenzierte und vergleichende Analysen zu befürworten sind.

Ein weiterer Vorteil, der in der vergleichenden Bildungsforschung gesehen wird, ist der Abbau von ethnozentristischen Vorstellungen. Ethnozentrismus ist die Tendenz, die eigene Gruppe oder Kultur als Maßstab vorauszusetzen, bspw. um andere Kulturen zu bewerten (Allemann-Ghionda 2004, S. 67). Ein möglicher Einfluss ethnozentristischer Sichtweisen findet sich bspw. in der PISA-2012-Studie beim Konzept „Einstellung gegenüber dem Lernen“. Dieses versucht die Einstellung unter anderem mit folgendem Item zu erfassen „Mich in der Schule anzustrengen wird mir den Zugang zum Studium ermöglichen.“ Hier zeigt sich die Verankerung der Forscher in der gymnasialen und universitären Bildung. Dieses Item erfasst die Einstellung gegenüber dem Lernen sehr einseitig und ignoriert die Strukturen des gegliederten Bildungssystems, welches Lernenden in nicht gymnasialen Klassen im Zugang zur universitären Bildung einschränkt. Ein stärkeres Bewusstsein über Schulartunterschiede in den Bildungszielen und Bildungschancen würde an dieser Stelle die Forschung und die Ergebnisinterpretation bereichern.

Vergleichende Bildungsforschung birgt zusammenfassend Gefahren der Vereinfachung und Zuschreibung. Auf der anderen Seite können damit die Reflexion über Bildung und Bildungskonzepte gestärkt und Erkenntnisse erweitert werden. Zu letzterem möchte diese Arbeit beitragen und den Blick für Unterschiede im Bildungssystem schärfen. Das Bildungssystem der Sekundarstufe I wurde in Kapitel 2 ausführlich beschrieben und es konnte deutlich gezeigt werden, dass sich die Schularten aus der historischen Entwicklung heraus in ihren Leitbildern, den Bildungszielen und -inhalten sowie in der Zusammensetzung der Schülergruppe unterscheiden. Zudem sind mit den spezifischen Zielen der Schularten auch Möglichkeiten und Grenzen für die inhaltliche Gestaltung des Unterrichts verbunden bspw. bezogen auf die angebotenen Unterrichtsfächer. Auch für die Lernenden ergeben sich Konsequenzen daraus, dass mit den Schularten je spezifische Bildungs- und Berufsperspektiven verbunden sind. Nicht zuletzt verteilen sich die Lehrenden und Lernenden nicht zufällig auf die Schularten, weshalb individuelle Merkmale mit den interessierenden Merkmalen der Schularten konfundiert sind. Es stellt sich daher auch die Frage, inwieweit die Schularten in ihren unterschiedlichen Ausgangsbedingungen überhaupt vergleichbar sind bzw. inwieweit anhand standardisierter Erhebungsinstrumente das schulische Lernen vergleichend erfasst werden kann.

Diese eher methodische Frage ist eine weitere Herausforderung, der sich die vergleichende Bildungsforschung zusätzlich zu den oben diskutierten inhaltlichen Vor- und Nachteilen des Vergleichs stellen muss. Denn neben Vereinfachungen und Pauschalisierungen durch Statistiken besteht auch die Gefahr, dass Unterschiede zwischen Gruppen durch ein unpassendes Messinstrument erst erzeugt werden (Allemann-Ghionda 2004). Diese Tatsachen erschweren aus methodischer Sicht die vergleichende Betrachtung der Schularten, da Unterschiede zwischen den Schularten neben tatsächlichen Unterschieden auch auf Artefakte zurückgeführt werden könnten, welche durch die Erhebung selbst hervorgerufen werden (van de Vijver 1998). Ein mögliches Beispiel ist das oben beschriebene Fragebogenitem, welches die Einstellung zum Lernen messen soll und allein schon durch seine Formulierung vermutlich Unterschiede zwischen den Schularten produziert. Denn Schülerinnen und Schüler, welche nicht in gymnasialen Klassen lernen, werden vermutlich bei gleicher Einstellung mit höherer Wahrscheinlichkeit die Aussage ablehnen, dass ihnen ein hoher Lerneinsatz den Zugang zum Studium ermöglicht als Lernende in gymnasialen Klassen. Somit würde auch bei gleicher Einstellung gegenüber dem

Lernen bei Schülerinnen und Schülern in gymnasialen Klassen ein höherer Einstellungswert gemessen als bei Lernenden nichtgymnasialer Klassen.

Solche nicht intendierten Varianzquellen, welche alternative Erklärungen für Gruppenunterschiede hervorrufen, werden in der vergleichenden Forschung auch unter dem Begriff des „Bias“ diskutiert (van de Vijver 1998). Bias bezieht sich auf störende Faktoren in der Erhebung, welche die gewünschte Äquivalenz, die Vergleichbarkeit von gemessenen Werten in verschiedenen Gruppen, beeinflussen können (He und van de Vijver 2012). Solche Einflüsse werden im Folgenden näher beschrieben.

4.1 Formen von Bias

Die verschiedenen Formen von Bias, welche die Äquivalenz von Messinstrumenten gefährden können, werden unter den Begriffen Konstruktbias, Methodenbias und Itembias subsumiert. Ein Konstruktbias liegt vor, wenn ein zu messendes Konstrukt, wie die Einstellung gegenüber der Schule, nicht identisch in den verschiedenen Vergleichsgruppen ist (van de Vijver 1998). Dies ist bspw. der Fall, wenn die verwendeten Items das Konstrukt in einer Gruppe nicht erfassen oder wenn nicht alle Aspekte, die das Konstrukt in den Gruppen umfasst, erhoben werden (van de Vijver und Poortinga 1997). In der kulturvergleichenden Forschung wird unter anderem das Empfinden von Glück (Happiness) als Beispiel für unterschiedliche Konstruktrepräsentationen zwischen Nordamerikanern und Ostasiaten genannt (He und van de Vijver 2012).¹³ Für Nordamerikaner erhöhen erreichte persönliche Ziele das Glücksempfinden. In ostasiatischen Ländern wird Glück eher als interpersonelle Verbundenheit definiert, welche sowohl positiv als auch negativ mit dem Glücksempfinden zusammenhängen kann (Uchida et al. 2004). In einem solchen Fall würden einheitliche Fragebogenitems die Konstrukte in den Vergleichsgruppen nicht äquivalent abdecken bzw. erheben. Auf den schulischen Kontext bezogen kann wiederum das obige Beispielitem aus der PISA-2012-Erhebung zur Einstellung gegenüber dem Lernen genannt werden. Die Relevanz schulischen Lernens für einen späteren Studienplatz ist systembedingt nicht für alle Lernenden der verschiedenen Schularten in gleicher Weise gegeben, weshalb die eingesetzten

13 Hier sei auf die Problematik des Kulturbegriffes, wie im einleitenden Teil dieses Kapitels beschrieben, verwiesen.

Items das Konstrukt Einstellung gegenüber dem Lernen in den Vergleichsgruppen nicht äquivalent erheben.

Von einem Bias können neben dem gesamten Konstrukt auch nur einzelne Items betroffen sein (*Itembias*). Dies ist beispielsweise der Fall, wenn verwendete Begriffe unterschiedlich besetzt sind (van de Vijver und Leung 1998). So könnte unter „Feedback“ im schulischen Kontext die Reaktion eines Lehrenden auf die Aussage eines Lernenden verstanden werden oder aber eine ausführliche (schriftliche) Rückmeldung zu Stärken und Schwächen bezogen auf einen längeren Zeitraum. Bei systematischen Unterschieden in der Wortdefinition zwischen den Vergleichsgruppen würden daher systematisch verschiedene Inhalte abgefragt.

Konstruktbias und Itembias beziehen sich auf die inhaltlichen Aspekte einer Fragebogenerhebung. Auf der anderen Seite kann auch die technische Umsetzung der Erhebung beeinträchtigt sein, was unter dem Begriff des Methodenbias diskutiert wird. Hier sind wiederum drei Formen zu unterscheiden. Zum einen kann die Stichprobe verzerrt sein (Sample Bias), beispielsweise wenn das Gruppierungsmerkmal (Lernende einer Schulart) stark mit anderen individuellen Faktoren konfundiert ist (bspw. die soziale Herkunft). Zum anderen kann sich der Erhebungsmodus systematisch zwischen den untersuchten Gruppen unterscheiden (Administration Bias), bspw. wenn die Zeiten zur Bearbeitung eines Leistungstests zwischen den Schularten verschieden wären. Als drittes kann der Prozess zur Beantwortung der Fragen verzerrt sein (Instrument Bias), bspw. wenn die Befragten im Umgang mit standardisierten Umfragen unterschiedlich erfahren sind. Auch Antworttendenzen wie die Tendenz, Fragen eher zuzustimmen (Akquieszenz; ARS), oder die Tendenz, eher mittlerer Antwortkategorien zu wählen (Midpoint Responding; MPR), können die Äquivalenz des Messinstrumentes im Sinne eines Instrument Bias minimieren (van de Vijver und Leung 2011).

In standardisierten Befragungen stellen Verzerrungen durch Antworttendenzen eine besondere Herausforderung dar, vor allem bei Items im Likert-Format (Bogner und Landrock 2015). Diese werden meist eingesetzt, wenn Einstellungen oder Verhaltensweisen erfragt werden, und sind dadurch auch in der Schul- und Unterrichtsforschung ein viel genutztes Instrument. Deshalb werden die Ursachen und Konsequenzen von Antworttendenzen im Folgenden ausführlicher beschrieben.

4.2 Antworttendenzen

Bei Items im Likert-Format können Befragte einer Aussage auf einer mehrstufigen Antwortskala mehr oder weniger zustimmen, bzw. diese mehr oder weniger ablehnen. Die Grundidee ist, dass Personen entsprechend ihrer Einstellung einzelne Items umso stärker ablehnen, je weiter ihre Einstellung von der formulierten Aussage entfernt ist (Wolf und Best 2010). Situationsbezogene Faktoren sowie dispositionelle Eigenschaften können die individuellen Antworten jedoch systematisch beeinflussen (Baumgartner und Steenkamp 2001; Weijters 2006). Systematische Einflüsse auf die Antworten von Befragten werden unter dem Begriff der Antworttendenzen erforscht. Situationsbezogene Faktoren, welche sich auf das Antwortverhalten von Befragten auswirken, sind der Erhebungsmodus (bspw. anonyme Befragung vs. Interview), das Format (bspw. vier vs. fünf Antwortmöglichkeiten) oder die erfragten Inhalte. Dispositionelle Erklärungen rücken hingegen Eigenschaften von Personen in den Vordergrund (Baumgartner und Steenkamp 2001). Hier setzt bspw. Forschung an, welche Antworttendenzen mit dem Alter, der Bildung oder der kognitiven Fähigkeit zu erklären versucht (siehe bspw. Lechner und Rammstedt 2015; Meisenberg und Williams 2008).

Antworttendenzen, die durch situationsbezogene Faktoren hervorgerufen werden, verändern die absoluten Messwerte der befragten Personen. In der Annahme, dass diese Faktoren in einer standardisierten Befragung alle Befragten in gleicher Weise beeinflussen, sollten Gruppenunterschiede nicht durch situationelle Faktoren verzerrt werden. Dispositionelle Einflüsse durch die Bildung oder die kognitiven Fähigkeiten stellen in der Schul- und Unterrichtsforschung jedoch eine Herausforderung dar. Lernende der verschiedenen Schularten unterscheiden sich bekanntermaßen im Durchschnitt in eben diesen Merkmalen (siehe Abschnitt 3.5.1). Antworttendenzen, welche durch Bildungsunterschiede oder die kognitive Fähigkeit bedingt sind, können sich somit in ihrem Ausmaß zwischen den Schularten unterscheiden und zu einer reduzierten Vergleichbarkeit der Gruppen führen. Die Entstehung von Antworttendenzen zu ergründen und deren Ausmaß einzuschätzen, ist daher in der Schul- und Unterrichtsforschung von besonderer Bedeutung.

Eine Erklärung für den Zusammenhang kognitiver Fähigkeiten und möglicher Antworttendenzen liefert das Modell von Cannell et al. (1981). Nach diesem

Modell durchlaufen Befragte im Idealfall einen Prozess aus fünf Schritten von Stimulus zur Antwortabgabe:

1. Verstehen der Frage
2. Abruf relevanter Informationen aus dem Gedächtnis
3. Beurteilung der gewonnenen Informationen hinsichtlich der Relevanz für die Frage
4. Beurteilung der gewonnenen Informationen hinsichtlich der Passung zur eigenen Meinung oder den eigenen Zielen
5. Antwortgabe

Kognitive Ermüdung kann jedoch dazu führen, dass Befragte den Antwortprozess vereinfachen und nicht alle Schritte vollständig durchlaufen (Krosnick 1991; Krosnick 2000). Bei der Beantwortung von Einstellungsfragen sind die Befragten beispielsweise aufgefordert einer Aussage zuzustimmen oder diese abzulehnen. Hierfür werden nach Erkenntnissen der Fragebogenforschung zunächst Gründe für eine Zustimmung abgerufen. Erst im Anschluss werden Gegenpositionen gesucht (Roßmann 2017). Befragte mit geringeren kognitiven Fähigkeiten könnten in diesem zweiten Schritt des Modells nach Cannell et al. (1981) schneller ermüden und den Prozess der Antwortfindung abbrechen, bevor Gegenpositionen gefunden wurden. In der Folge werde Items unabhängig vom Inhalt schneller zugestimmt. Diese inhaltsunabhängige Tendenz, Fragen zuzustimmen, wird auch als Akquieszenz (acquiescence response style; ARS) bezeichnet (Baumgartner und Steenkamp 2001).

Eine andere Strategie, die kognitiven Anforderungen zu minimieren, betrifft die Beurteilung der gewonnenen Informationen im dritten und vierten Schritt des Modells nach Cannell et al. (1981). Werden die abgerufenen Informationen nicht ausreichend beurteilt, führt eine geringe Differenzierung zur Tendenz, mittlere Antwortkategorien zu wählen (Krosnick 1991; Krosnick 2000). Diese Antworttendenz wird in der Forschung auch unter dem Begriff des Midpoint Responding (MPR) beschrieben (Baumgartner und Steenkamp 2001).

Je schwieriger die Fragen und je geringer die kognitiven Fähigkeiten der Befragten, desto wahrscheinlicher ist es, dass Befragte vom optimalen Prozess abweichen und auf alternative Strategien wie ARS und MPR zurückgreifen (Roßmann 2017). Da sich die Lernenden der verschiedenen Schularten in ihren kognitiven Fähigkeiten unterscheiden, ist zu vermuten, dass die aggregierten Antworten in den verschiedenen Schularten unterschiedlich stark durch Antworttendenzen wie ARS und MPR belastet sind.

4.3 Implikationen für den Vergleich von Schularten

Der einleitende Teil dieses Kapitels hat die Gefahren und Potentiale der vergleichenden Forschung offengelegt. Besonders die statistischen Vereinfachungen beim Vergleich der Schularten können zu pauschalisierten Schlussfolgerungen führen. Für die Verfasser wissenschaftlicher Vergleichsforschung wie auch für die Rezipienten der vergleichenden Literatur gilt es sich dieser Gefahren bewusst zu bleiben.

Aus den vorangegangenen Abschnitten wird zudem deutlich, dass aus methodischer Sicht die Vergleichbarkeit von Fragebogendaten zwischen den Schularten eingeschränkt sein kann – zum einen durch abweichende Konstruktrepräsentationen in den Vergleichsgruppen oder unterschiedliche Verständniszuschreibung zu einzelnen Items (Abschnitt 4.1). Zum anderen können zusätzlich schulartspezifische Antworttendenzen, welche durch Unterschiede in den kognitiven Fähigkeiten der Lernenden verschiedener Schularten bedingt sein können, die Daten systematisch zwischen den Gruppen verzerren (Abschnitt 4.2). Bevor Gruppenvergleiche basierend auf Fragebogendaten angestellt werden, gilt es daher zu prüfen, ob die eingesetzten Instrumente die Daten äquivalent in den zu vergleichenden Gruppen messen (Davidov 2008). Zusätzlich muss die Auswirkung von Antworttendenzen geprüft werden (Liu et al. 2017).

Für die Überprüfung der Äquivalenz von Messinstrumenten hat sich in der empirisch quantitativen Forschung der Ansatz der Invarianzanalyse durchgesetzt (Davidov 2008; van de Vijver und Leung 1998). Der Begriff Invarianz bezeichnet allgemein die Unveränderlichkeit von Größen (Wahrig und Wahrig-Burfeind 2003). Im Kontext der Fragebogenforschung bedeutet Invarianz, dass die Annahmen über die Zusammenhänge von Variablen zwischen den zu vergleichenden Gruppen unveränderlich, also invariant sein müssen, damit die

zugehörigen Konstrukte valide zwischen Gruppen verglichen werden können. Die Invarianz wird hierbei in drei hierarchischen Stufen getestet (Steenkamp und Baumgartner 1998) und hat in ihren verschiedenen Stufen unterschiedliche Konsequenzen für die Vergleichbarkeit der Daten. Die Invarianzanalyse und mögliche Konsequenzen für die Vergleichbarkeit werden im Methodenteil näher erklärt (Abschnitte 6.4.1 und 6.4.2).

Antworttendenzen können als eine Form des Methodenbias die Invarianz von Messinstrumenten reduzieren (siehe bspw. Cheung und Rensvold 2000; Grimm und Church 1999). Nicht immer kann eine Invarianzanalyse jedoch die Verzerrung der Daten durch Antworttendenzen aufdecken. So zeigen Simulationsstudien, dass nur bei extremen Verzerrungen mangelnde Invarianz durch Antworttendenzen aufgedeckt werden kann (Liu et al. 2017). Es empfiehlt sich daher einzelne Formen der Antworttendenzen zusätzlich zu prüfen (siehe im Methodenteil Abschnitt 6.4.3).

Soweit aus bisherigen Studien, welche Schulartunterschiede in Deutschland untersuchen ersichtlich (siehe Abschnitt 3.5), wurde in keiner der aufgeführten Studien die Vergleichbarkeit der Daten zwischen den Schularten geprüft. Dies ist ein erhebliches Forschungsdesiderat. Die bisherigen Ergebnisse könnten in ihrer Validität eingeschränkt sein, weshalb bisherige Erkenntnisse über Schulartunterschiede an Glaubhaftigkeit einbüßen, solange die Vergleichbarkeit der Daten nicht geprüft ist. Dieser Forschungslücke nimmt sich die hier vorliegende Arbeit an, um die Bedeutung der bisherigen Ergebnisse zu Schulartunterschieden zu prüfen und den Erkenntnisstand zu erweitern. Die Forschungsfragen dieser Arbeit sind in Kapitel 5 näher spezifiziert.

5 Herleitung der empirischen Fragestellungen

Die vorliegende Arbeit befasst sich mit der Bedeutung des gegliederten Bildungssystems der Sekundarstufe I für die Prozesse des Unterrichts. Hierfür wurden zunächst in Kapitel 2 die Entwicklung des Bildungssystems nachgezeichnet und die resultierenden Unterschiede in der aktuellen Ausgestaltung zwischen den Schularten herausgearbeitet. Besonders in den Bildungszielen und -inhalten, den Leitbildern und in den Selektionsprozessen sind deutliche Unterschiede erkennbar.

Kapitel 3 schließt mit Überlegungen an, wie sich diese systembedingten Unterschiede auf den Unterricht der Lernenden auswirken können. Im Rückgriff auf didaktische Theorien wurde ein erweitertes Didaktisches Dreieck hergeleitet, welches die Vermittlungsmechanismen über den Lehrenden, den Gegenstand sowie die Lernenden verdeutlicht:

- Die lehramtsspezifische Ausbildung zeichnet sich durch unterschiedliche Lerngelegenheiten aus. An den Universitäten manifestieren sich diese Unterschiede in den verschiedenen Schwerpunktlegungen im Fachwissen, im fachdidaktischen Wissen sowie im didaktischen Wissen für die Lehramtsstudiengänge. In der zweiten Ausbildungsphase an den Studienseminaren und Schulen differenzieren sich die schulartspezifischen Lerngelegenheiten weiter aus. Damit gehen Unterschiede in den Kompetenzen und den Überzeugungen der Lehrenden einher, welche wiederum in direktem Zusammenhang mit dem professionellen Handeln stehen (siehe Abschnitt 3.2).
- Besonders für die inhaltsorientierten Lehrpläne zeigen sich Unterschiede zwischen den Schularten. Diese schulartspezifischen Lehrpläne und Curricula werden auf Schul- und Klassenebene entsprechend den Lernzielen der jeweiligen Schulart konkretisiert. Das intendierte Curriculum gibt damit den Handlungsspielraum in der inhaltlichen und zeitlichen Gestaltung des Unterrichts vor und erhöht die Wahrscheinlichkeit, dass bestimmte Inhalte bzw. Gegenstände gelehrt und andere ausgeschlossen werden (siehe Abschnitt 3.3).
- Selektions- und Segregationsprozesse führen zu einer bestimmten Zusammensetzung der Schülerschaft an den verschiedenen Schularten. Kompositionseffekte verdeutlichen die Bedeutung, welche der Schülergruppe für die

Entwicklung des Einzelnen zukommt. Aus den theoretisch ausgearbeiteten Wirkmechanismen lassen sich Hypothesen zu Effekten der Schülergruppe auf das Lehren und Lernen ableiten. Vor allem ist zu erwarten, dass einerseits das kognitive Lernumfeld und die Disziplin in der Klasse von der Zusammensetzung der Klasse abhängen (instrumenteller Peer-Effekt) und damit auch von der Schulart, da Komposition und Schulart hoch korreliert sind. Andererseits ist zu erwarten, dass die Lehrpersonen entsprechend ihrer Erwartungen gegenüber der Schülergruppe unterschiedliche Anforderungen stellen (siehe Abschnitt 3.4).

Die hier vorliegende Arbeit geht in den theoretischen Überlegungen damit einen Schritt weiter, als es bislang in der Diskussion um Lern- und Entwicklungsmilieus der Fall war (bspw. Baumert et al. 2006b). Didaktische Traditionen, das Curriculum sowie die Schülergruppe wurden dort zwar als theoretische Erklärungsfaktoren für die unterschiedlichen Entwicklungstendenzen der Lernenden verschiedener Schularten genannt. Überlegungen, wie diese Faktoren mit Prozessen der Unterrichtsebene zusammenhängen, wurden hingegen weitestgehend ausgeblendet.

Der Überblick zum Forschungsstand (Abschnitt 3.5) arbeitet deutliche Schulartunterschiede in den konstituierenden Komponenten des Unterrichts – den Lehrenden, den Inhalten und der Schülergruppe – sowie in den Unterrichtsprozessen – konkret den generischen Faktoren des Unterrichts und den eingesetzten Methoden – heraus. An Analysen, die Unterschiede in den Unterrichtsprozessen zwischen den Schularten mit Merkmalen der Lehrenden, der Schülergruppe oder mit Inhalten des Unterrichts erklären, scheint es hingegen bislang zu mangeln. Zudem wird aus dem Forschungsüberblick deutlich, dass die bisherigen Ergebnisse einige methodische Schwächen aufweisen.

Zum einen werden bei den deskriptiven Vergleichen meist keine Signifikanztests berichtet, welche die Varianz innerhalb und zwischen den Schularten berücksichtigen. Daher ist nicht auszuschließen, dass beschriebene Unterschiede teils dem Zufall unterliegen. Zudem werden nur selten die Schularten Realschule, Schulen mit mehreren Bildungsgängen und integrierte Gesamtschulen in der Forschung separat analysiert oder beschrieben. Besonders für diese Schularten fehlen daher Informationen für deren umfassende Darstellung.

Darüber hinaus lässt sich an der bisherigen Forschung kritisieren, dass wichtige Vorannahmen bezüglich der Vergleichbarkeit der verwendeten Indikatoren für Schulartvergleiche nicht geprüft wurden. Kapitel 4 beschreibt potentielle Einschränkungen in der Vergleichbarkeit der Schularten und argumentiert, dass in den Fragebogenerhebungen der Schul- und Unterrichtsforschung durch die Unterschiede in den Vergleichsgruppen und durch die systembedingten Unterschiede zwischen den Schularten die Bedingungen für vergleichende Analysen verletzt sein können. Konkret wird angenommen, dass einzelne Konstrukte keine ausreichende Invarianz für den Vergleich von Mittelwerten und Zusammenhängen zwischen den Schularten aufweisen könnten und die Fragebogendaten durch Antworttendenzen verzerrt sein könnten. Dies würde bedeuten, dass möglicherweise die bisher gefundenen Unterschiede zwischen den Schularten in der Ausprägung erfragter Merkmale lediglich auf Artefakte zurückzuführen sind.

Aus den theoretischen Überlegungen zu Auswirkungen des gegliederten Bildungssystems der Sekundarstufe I auf die Unterrichtsprozesse in verschiedenen Schularten und den aufgezeigten Forschungsdesideraten können vier Fragekomplexe abgeleitet werden, welche diese Arbeit empirisch bearbeitet. Übergeordnetes Ziel ist es, Unterrichtsprozesse und damit zusammenhängende Faktoren vergleichend zwischen den Schularten zu betrachten. Um die Validität dieses Vergleichs sicherzustellen, wird zuerst die Vergleichbarkeit der verwendeten Skalen zwischen den Schularten geprüft (Fragestellung 1). Darauf aufbauend interessiert die Replizierbarkeit bisheriger Forschungsbefunde unter Kontrolle von potentiellen Antworttendenzen und unter Einbezug der Merkmalsvarianz zwischen und innerhalb der Schularten (Fragestellung 2). Anschließend wird der Frage nachgegangen, welche Faktoren auf Seiten der Lehrenden, der Schülergruppe sowie der Unterrichtsinhalte Unterschiede in den Unterrichtsprozessen zwischen den Schularten erklären können (Fragestellung 3). Abschließend wird untersucht, wie die Unterrichtsprozesse innerhalb der einzelnen Schularten mit der Leistung zusammenhängen, bzw. ob sich die Stärke oder sogar die Richtung dieser Zusammenhänge zwischen den Schularten unterscheiden (Fragestellung 4).

1. Inwieweit sind die in dieser Arbeit für den Vergleich von Schularten herangezogenen latenten Konstrukte vergleichbar zwischen den Schularten?

Diese Frage bezieht sich zunächst darauf, ob die latenten Konstrukte äquivalent zwischen den Gruppen erfasst und somit die Ergebnisse zuverlässig interpretiert werden können. Der Vergleich von Zusammenhängen setzt konfigurale und metrische Invarianz voraus. Der Vergleich von Mittelwerten erfordert zudem skalare Invarianz (Davidov 2008; siehe auch Abschnitt 4.1). Ziel der Arbeit ist es nicht nur statistisch die Vergleichbarkeit zu prüfen, sondern auch inhaltlich mögliche Ursachen für mangelnde Vergleichbarkeit zu ergründen. Hierfür werden bei mangelnder konfiguraler Invarianz die konkreten Fragebogenitems inhaltlich analysiert. Konkret werden folgende Fragen untersucht:

- a. Welche latenten Konstrukte sind konfigural und metrisch invariant zwischen Schularten und erlauben daher Vergleiche von Zusammenhängen?
- b. Welche latenten Konstrukte sind darüber hinaus skalar invariant zwischen Schularten und erlauben daher Mittelwertvergleiche?
- c. Lassen sich durch die Itemformulierungen mögliche Unterschiede zwischen den Schularten hinsichtlich der latenten Konstrukte erklären?

Die Daten können zusätzlich durch Antworttendenzen verzerrt sein, was die Vergleichbarkeit ebenfalls einschränken kann. Besonders, wenn sich das Ausmaß der Antworttendenzen zwischen den Schularten unterscheidet, schränken systematische Verzerrungen die Zuverlässigkeit der Ergebnisse ein. Da sich die Lernenden in ihren kognitiven Fähigkeiten zwischen den Schularten im Mittel unterscheiden und gerade mit sinkender kognitiver Fähigkeit die Wahrscheinlichkeit für Antworttendenzen steigt (Krosnick 2000, siehe auch Abschnitt 4.2), ist von systematischen Verzerrungen auszugehen. Insofern wäre zu erwarten, dass Lernende an Schulen mit mehreren Bildungsgängen und an integrierten Gesamtschulen im Mittel stärkere Antworttendenzen aufweisen als Lernende an Realschulen und insbesondere an Gymnasien. Die Lehrenden sollten hingegen über die Schularten hinweg ähnliche Ausprägungen in den Antworttendenzen aufzeigen. Um die bisherigen Ergebnisse in ihrer Zuverlässigkeit besser interpretieren zu können, interessiert zudem, wie stark Ergebnisse ohne Kontrolle von Antworttendenzen von Ergebnissen unter Kontrolle von Antworttendenzen abweichen.

- d. Unterscheiden sich Antworttendenzen von Lernenden systematisch zwischen den Schularten?
 - e. Unterscheiden sich die Antworttendenzen von Lehrenden systematisch zwischen den Schularten?
 - f. Wie stark unterscheiden sich die Ergebnisse von Mittelwertvergleichen für die hier untersuchten Skalen mit und ohne Kontrolle der Antworttendenzen?
2. *Welche Befunde aus der Forschung zu Schulartunterschieden im Lehren und Lernen können unter Kontrolle von Antworttendenzen und unter Berücksichtigung der Varianz der Merkmale zwischen und innerhalb der Schularten (Eta Quadrat) repliziert werden?*

Die Ergebnisse aus anderen Studien zu Schulartvergleichen (siehe Abschnitt 3.5) können wie folgt zusammengefasst werden: Relativ übereinstimmend finden sich über verschiedene Schulfächer hinweg Unterschiede zwischen den Schularten in generischen Faktoren des Unterrichts und in den pädagogischen Orientierungen von Lehrpersonen. In einzelnen Studien können zudem Unterschiede im eingesetzten Aufgabenmaterial sowie in den Unterrichtszielen der Lehrenden identifiziert werden. Der Unterricht an Gymnasien zeichnet sich durch ein höheres Potential zur kognitiven Aktivierung aus. Dem Reflektieren, Modellieren und Argumentieren wird von Lehrenden an Gymnasien eine größere Bedeutung zugesprochen als von Lehrenden, die an den anderen Schularten unterrichten. Auf die Bedürfnisse der Lernenden gehen die Lehrenden an Gymnasien hingegen weniger ein als an allen anderen Schulformen. Auch die Unterrichtsmethoden sind dort häufiger lehrerzentriert als an den anderen Schularten. Für die Hauptschulen wurde das Gegenteil beobachtet: Das Potential zur kognitiven Aktivierung ist hier vergleichsweise gering und Lehrende fokussieren eher darauf, Grundlagen und Routinen einzuschleifen; die konstruktive Unterstützung der Lernenden durch die Lehrenden ist hingegen hoch. Letzteres zeigt sich auch in einem vergleichsweise häufigeren Einsatz schülerorientierter Unterrichtsmethoden. Realschule, Schulen mit mehreren Bildungsgängen und integrierte Gesamtschulen befinden hinsichtlich der Ausprägung dieser Merkmale überwiegend zwischen den Hauptschulen und den Gymnasien. In vielen Studien werden diese Schularten jedoch nicht separat beschrieben, weshalb hier kein eindeutiges Bild vorherrscht.

Da bei den deskriptiven Vergleichen meist keine Signifikanztests berichtet werden, welche die Varianz innerhalb und zwischen den Schularten berücksichtigen, könnten sich die Schularten in geringerem Umfang voneinander unterscheiden, als die bisherigen Analysen suggerieren. Auch die Kontrolle von Antworttendenzen könnte sich auf die Unterschiede zwischen den Schularten auswirken. Da sowohl die Zustimmungstendenz (ARS) als auch die Tendenz, mittlere Antworten zu bevorzugen (MPR), mit sinkender kognitiver Fähigkeit steigt, ist jedoch schwer prognostizierbar, ob sich die Schulartunterschiede einzelner Merkmale unter Kontrolle von Antworttendenzen verstärken oder abschwächen. Die deskriptiven Mittelwertvergleiche sind daher eher induktiv als hypothesengeleitet. Die Ergebnisse werden hinsichtlich dreier leitender Fragen interpretiert:

- a. Welche Unterschiede zwischen den Schularten in den Unterrichtsprozessen und damit zusammenhängenden Faktoren lassen sich auf der Basis der hier verwendeten Daten mit Kontrolle von Antworttendenzen und Varianzen in den Merkmalsstrukturen identifizieren?
 - b. Inwieweit stimmen die Ergebnisse mit jenen, die in früheren Studien beobachtet wurden, überein bzw. weichen davon ab?
 - c. Welches sind die zentralen Merkmale, in denen sich die Schularten voneinander unterscheiden?
3. *Können Merkmale von Lehrenden, der Schülergruppe sowie der Unterrichtsinhalte Unterschiede in den Unterrichtsprozessen zwischen den Schularten erklären?*

Eine Kernthese dieser Arbeit ist, dass durch das gegliederte Bildungssystem die Gestaltungsmöglichkeiten des Lehrens und Lernens durch die schulartspezifischen Leitbilder, die Lehrerausbildung, die Lehrpläne sowie die Selektions- und Segregationsprozesse beeinflusst sind. Dies impliziert, dass Unterschiede zwischen den Schularten in den Unterrichtsprozessen über bestimmte Mechanismen vermittelt werden: die Überzeugungen und die Ausbildung der Lehrenden, die Zusammensetzung der Schülergruppe und die Inhalte des Unterrichts (siehe Kapitel 3). Diese These gilt es empirisch zu prüfen. Deshalb soll für alle Unterrichtsprozessmerkmale, welche bedeutende Schulartunterschiede aufweisen, geprüft werden, ob die Effekte der Schulart auf die Unterrichtsprozesse

durch Überzeugungen der Lehrenden, die Zusammensetzung der Schülergruppe und/oder Inhalte des Unterrichts vermittelt sind.

- a. Für welche Merkmale der Lehrenden, der Schülergruppe und der Unterrichtsinhalte finden sich Zusammenhänge mit den untersuchten Unterrichtsprozessmerkmalen?
 - b. Vermitteln diese Merkmale der Lehrenden, der Schülergruppe und der Unterrichtsinhalte die Effekte der Schulart auf die Unterrichtsprozessmerkmale?
4. *Unterscheiden sich die Zusammenhänge zwischen den Unterrichtsprozessmerkmalen und der Leistung in Mathematik in der 10. Klasse unter Kontrolle der Vorleistung in der 9. Klasse zwischen den Schularten?*

In der Schuleffektivitätsforschung werden anknüpfend an konstruktivistische und motivationspsychologische Perspektiven auf unterrichtliches Lernen Unterrichtsprozessmerkmale auf ihren Zusammenhang mit der Entwicklung der Lernenden untersucht. Die Annahme ist, dass lernwirksamer und motivationsförderlicher Unterricht sich durch die kognitiv aktivierende Behandlung fachlich zentraler Inhalte, durch eine effektive Klassenführung mit wenigen Unterrichtsstörungen und durch ein positiv geprägtes Unterrichtsklima auszeichnet (Lipowsky 2015). Auch bei den eingesetzten Unterrichtsmethoden wird versucht zu ermitteln, welche Methoden – eher lehrerzentrierte oder eher schülerorientierte – bezüglich der Entwicklung der Lernenden überlegen sind (siehe auch Abschnitt 3.1). Formatives Assessment dient nicht direkt der Vermittlung oder Aneignung von Inhalten, sondern ist eher als übergreifende Maßnahme zur Lehr- und Lernplanung zu sehen. Aber auch hier werden Zusammenhänge mit der motivationalen und kognitiven Entwicklung der Lernenden angenommen und untersucht (einen Überblick gibt bspw. Hattie 2009). Es zeigt sich jedoch, dass die Befundlage nicht eindeutig ist. Besonders in groß angelegten Leistungsvergleichstests, welche die Vielfalt der Schulen und Schularten in Deutschland repräsentativ abdecken, zeigen sich die postulierten Zusammenhänge der Unterrichtsprozessmerkmale mit Leistung und Leistungsentwicklung häufig nicht. Teilweise werden sogar konträre Zusammenhänge gefunden, indem formatives Assessment oder konstruktive Unterstützung negativ mit der Entwicklung der Lernenden zusammenhängen (bspw. Kuger et al. 2017).

Eine mögliche Erklärung für die inkonsistenten Befunde könnte sein, dass die Unterrichtsprozessmerkmale, welche in Deutschland in Gymnasien und Realschulen signifikante Zusammenhänge mit Schulleistungen zeigen, für alle anderen Schularten nicht gleichermaßen relevant sind. Daher werden abschließend die Zusammenhänge zwischen Unterrichtsprozessmerkmalen, welche sich in ihren Ausprägungen bedeutend zwischen den Schularten unterscheiden, hinsichtlich ihres Zusammenhangs mit der Leistung analysiert, getrennt nach Schulart. Leitend für die Interpretation sind folgende Fragen:

- a. Finden sich signifikante Zusammenhänge zwischen den Unterrichtsprozessmerkmalen und der Leistung in Mathematik in der 10. Klasse unter Kontrolle der Vorleistung in der 9. Klasse?
- b. Unterscheiden sich diese Zusammenhänge hinsichtlich ihrer Stärke oder sogar hinsichtlich ihrer Richtung zwischen den Schularten?

6 Methode

Die nationale PISA-2012-Studie ist die Datenbasis für den empirischen Teil dieser Arbeit. Die PISA-Studien untersuchen die Kompetenzen von 15-jährigen Schülerinnen und Schülern in mehreren Kompetenzbereichen. In der Erhebung 2012 wurde schwerpunktmäßig die mathematische Kompetenz der Lernenden untersucht. Zusätzlich zu den Kompetenztests erhebt die PISA-Studie Informationen über die Lernenden, die Lehrenden, den Unterricht und die Schule (Hintergrundfragebögen; Prenzel et al. 2013). Mit der PISA-Studie kann somit der Unterricht der Lernenden umfassend beschrieben werden und es lassen sich Zusammenhänge von Unterrichtsprozessen mit Merkmalen der Schülergruppe, der Lehrenden und der Unterrichtsinhalte untersuchen.

Nationale Erweiterungen in der deutschen PISA-Studie erlauben die Beschreibung des Unterrichts auf Klassenebene, da zusätzlich zu den 15-jährigen Schülerinnen und Schülern vollständige Klassen der neunten Jahrgangsstufe im Schuljahr 2012 erhoben wurden. Als weitere Besonderheit der nationalen PISA-2012-Studie nahmen die gleichen Lernenden zudem am Ländervergleich teil (vgl. Hecht et al. 2013). Die Lernenden bearbeiteten daher am ersten Testtag die PISA-Aufgaben (im Folgenden PISA-Test genannt) und am zweiten Testtag den Test des Ländervergleichs zur Überprüfung der erreichten Bildungsstandards (im Folgenden BiSta-Test genannt). Somit liegen als weitere Besonderheit der nationalen Erweiterungen die Leistungsdaten der Lernenden für unterschiedliche Leistungstests in Mathematik vor. Nach ca. einem Jahr Abstand wurden die Teilnehmenden erneut gebeten, die Testaufgaben und Hintergrundfragebögen zu bearbeiten (Heine et al. 2017). Somit sind mit der nationalen PISA-2012-Studie durch die Erweiterungen des nationalen Designs sogar Längsschnittanalysen möglich.

Die realisierte Stichprobe der nationalen PISA-2012-Studie sowie die verwendeten Indikatoren zur Beantwortung der Fragestellungen dieser Arbeit werden im Folgenden näher erläutert. Die anschließenden Abschnitte beschreiben die Analysemodelle zur Beantwortung der Fragestellungen.

6.1 Stichprobe

Die Hintergrundfragebögen für die Lernenden wurden in PISA 2012 in einem Rotationsdesign mit drei überlappenden Booklets ausgegeben, d.h., jede Frage war in zwei der drei Booklets vorhanden und wurde von etwa zwei Dritteln der Lernenden beantwortet (Sälzer und Prenzel 2013). Die Zuteilung der Booklets erfolgte innerhalb der Schulen und Klassen per Zufall. Die Fragen zu den Unterrichtsprozessen waren in den gleichen Booklets (z. B. Booklet 1 und 2) untergebracht, die Fragen zu den Inhalten des Mathematikunterrichts in einem überlappenden und dem dritten Booklet (z. B. Booklet 2 und 3). Die Stichprobe variiert somit leicht, je nach untersuchtem Merkmal. Werden Merkmale der Unterrichtsprozesse gemeinsam mit Merkmalen der Unterrichtsinhalte analysiert, liegen aufgrund des Rotationsdesigns vollständige Antworten von ca. einem Drittel der Lernenden vor (Kuger et al. 2017).

Durch das Längsschnittdesign wird die Stichprobe zudem eingeschränkt. Für die hier vorliegenden Analysen werden nur Daten von Lernenden berücksichtigt, die an beiden Messzeitpunkten die Tests bearbeitet haben und zwischen den Erhebungszeitpunkten nicht die Klasse gewechselt haben. Zudem werden nur Lernende berücksichtigt, die eine der folgenden Schularten besuchen: Gymnasium, Realschule, integrierte Gesamtschulen und Schulen mit mehreren Bildungsgängen. Die 2596 Schülerinnen (51,1 %) und Schüler (48,9 %) dieses Subsamples gingen in 220 Klassen, welche sich auf die vier untersuchten Schularten wie folgt verteilen: 30 Klassen gehören Schulen mit mehreren Bildungsgängen an, 66 Klassen waren an Realschulen und 21 Klassen an integrierten Gesamtschulen. Gymnasien stellen mit einer Anzahl von 103 Klassen den größten Anteil. Die unterschiedlichen Anteile, welche teils durch systematische Ausfälle bedingt sind, werden in den Analysen durch entsprechende Gewichte berücksichtigt. Hauptschulen konnten nicht berücksichtigt werden, da diese häufig keine verpflichtende zehnte Jahrgangsstufe anbieten. Eine ausführliche Beschreibung der realisierten Stichprobe und der systematischen Ausfälle geben Heine et al. (2017).

Zusätzlich zu den Angaben der Lernenden werden die Aussagen von Mathematiklehrkräften zum Unterricht einbezogen. Berücksichtigt werden Mathematiklehrkräfte, die eindeutig einer Schule zugeordnet werden können. Für die Analysen in Fragestellung 1 wird hierfür auf das Gesamtsample der erhobenen und zuordenbaren Mathematiklehrerinnen (5,4 %) und Mathematiklehrer (96,6 %)

zurückgegriffen, da hier die Validität der Messinstrumente geprüft wird. Die 260 Lehrenden unterrichten an Gymnasien (N=105), Realschulen (N=77), integrierten Gesamtschulen (N=26) und Schulen mit mehreren Bildungsgängen (N=52). Die Fragestellungen 2 bis 4 beziehen sich konkret auf den Unterricht der Lernenden. Hierfür werden nur die Angaben der 220 Klassenlehrerinnen und -lehrer im Fach Mathematik einbezogen.

6.2 Instrumente

Die Fragebogeninstrumente und Leistungstests werden hier nur kurz beschrieben, da diese an anderer Stelle gut dokumentiert und die Informationen frei im Web verfügbar sind. Das PISA 2012 Skalenhandbuch (Mang et al. 2018) enthält die Itemformulierungen und Antwortskalen. Zudem sind deskriptive Statistiken und Informationen zur Skalengüte enthalten. Grundsätzlich sind in die Analysen dieser Arbeit nur Skalen eingegangen, welche auf nationaler Ebene eine ausreichende Messqualität (Cronbachs $\alpha \geq 0,7$) aufweisen. Weitere Informationen zu den theoretischen Konstrukten und der Skalierung fasst der PISA 2012 Technical Report (OECD 2014) zusammen. In den Dokumentationen liegen jedoch keine Informationen zur Invarianz der Konstrukte zwischen den Schularten vor, weshalb in dieser Arbeit die Invarianz der Konstrukte zur Beschreibung von Unterrichtsprozessen, des normativen Lernumfeldes und der Überzeugungen von Lehrenden auf verschiedene Stufen der Invarianz getestet werden (siehe Kapitel 4). Die Ergebnisse sind in Abschnitt 7.1.1 festgehalten.

Die untenstehenden Überblickstabellen fassen die in den Analysen verwendeten Skalen der Lernenden (Tabelle 4) und der Lehrenden (Tabelle 5) nach Themenkomplexen zusammen und listen die Konstruktamen sowie die Kurzbezeichnungen der Konstrukte auf. Somit sind die Informationen im Skalenhandbuch und Technical Report leicht zu finden. Zusätzlich wird in untenstehenden Tabellen angegeben, welche Konstrukte auf Likert-Items basieren. Dies ist hier relevant, da für diese Konstrukte die Auswirkungen von Antworttendenzen analysiert (Abschnitt 7.1.2) und in den Analysen zu Fragestellung 2 bis 4 (Abschnitte 7.2 bis 7.4) kontrolliert werden. Alle in dieser Arbeit einbezogenen Merkmale wurden in der 9. Klasse im Jahr 2012 erhoben. Einzige Ausnahme bilden die Leistungstests, welche in der 9. Klasse im Jahr 2012 und in der 10. Klasse im Jahr 2013 erhoben und in dieser Arbeit für Längsschnittanalysen einbezogen werden.

6.2.1 Skalen, Indizes und Leistungsdaten der Lernenden

In PISA 2012 wurden für die Beschreibung der Unterrichtsprozesse aus Sicht der Lernenden zwei verschiedene Zugänge gewählt. Die Lernenden beschreiben 1.) die Unterrichtsmethoden und 2.) die generischen Faktoren des Unterrichts. Alle Konstrukte, welche Unterrichtsprozesse beschreiben, werden mit Likert-Items gemessen. Die Inhalte des Mathematikunterrichts beschreiben die Lernenden, indem sie berichten, wie vertraut sie mit bestimmten Aufgaben und Konzepten sind und in welchem Umfang sie in Mathematik unterrichtet werden (siehe Tabelle 4).

Das kognitive Lernumfeld wird über die mittlere Mathematikleistung der Lernenden im PISA-Test und im BiSta-Test gemessen. Als eine Besonderheit liegen mit der PISA-2012-Studie für Deutschland zwei unterschiedliche Leistungstests für das gleiche Unterrichtsfach vor, welche jeweils in der 9. Klasse im Jahr 2012 und in der 10. Klasse im Jahr 2013 erhoben wurden. Der Hauptunterschied zwischen diesen beiden Testverfahren ist die curriculare Validität. Während der PISA-Test so gestaltet ist, dass er curriculumsunabhängig die Anwendung gelernter Inhalte, Fähig- und Fertigkeiten, also „mathematical literacy“, erhebt (Prenzel et al. 2013), berücksichtigt der Test des Ländervergleichs, der unter Mithilfe von Lehrenden und Fachdidaktikern erstellt wird, eine möglichst breite Spanne der Curricula an deutschen Schulen (Roppelt et al. 2013). Zusätzlich zu den fachspezifischen Leistungstests liegen mit dem Berliner Test zur Erfassung fluider und kristalliner Intelligenz für die 8. bis 10. Jahrgangsstufe (BEFKI 8-10) Informationen über den Intelligenzquotienten (IQ) der Lernenden vor. Der BEFKI 8-10 ist ein theoretisch fundierter Intelligenztest zur Erfassung allgemeiner kognitiver Fähigkeiten (Wilhelm et al. 2014).

Tabelle 4. Skalen, Indizes und Leistungsdaten der Lernenden aus PISA 2012 (Mang et al. 2018).

Themenkomplex	Konstruktname	Bezeichnung	Likert-Format
Unterrichtsprozesse	Unterrichtsmethoden	Formatives Assessment	✓
		Schülerorientierte Unterrichtsmethoden	✓
		Lehrerorientierter Unterricht	✓
	Generische Faktoren des Unterrichts	Konstruktive Unterstützung durch die Lehrperson im Mathematikunterricht	✓
		Kognitive Aktivierung im Mathematikunterricht	✓
		Klassenführung im Mathematikunterricht	✓
		Disziplin im Klassenzimmer in Mathematik	✓
Unterrichtsinhalt		Häufigkeit mathematischer Anwendungsaufgaben	
		Häufigkeit innermathematischer Aufgaben	
		Vertrautheit mit mathematischen Konzepten	
		Anzahl der Minuten in Mathematik	
Lernumfeld	Kognitiv	PISA Leistungstest in Mathematik 2012	standT1_mpvX_w1
		PISA Leistungstest in Mathematik 2013	standT2_mpvX_w1
		BiSta Leistungstest in Mathematik 2012	BISTA_pv1GL
		BiSta Leistungstest in Mathematik 2013	BISTA_mgl1_v0X
		IQ (BEFKI 8-10)	pv.X.gff
	Sozio-ökonomisch	Index der höchsten beruflichen Stellung der Eltern	HISEI
		Index des höchsten Bildungsabschlusses der Eltern	HISCED
	Normatives	Schüler-Lehrer-Beziehung	✓
		Allgemeine Einstellung zur Schule	✓
		Allgemeine Einstellung zur Lernaktivitäten	✓
		Arbeitsmoral	✓
		Mathematikbezogene instrumentelle Motivation	✓
		Freude und Interesse an Mathematik	✓

Ein Häkchen (✓) in der letzten Spalte gibt an, dass das Konstrukt mit Likert-Items gemessen wird.

Das sozioökonomische Lernumfeld wird über die Klassenmittelwerte zweier Merkmale beschrieben: Der sozioökonomische Hintergrund der Lernenden wurde durch den „International Socio-Economic Index of Occupational Status“ (ISEI; Ganzeboom, Graaf, Treiman & Leeuw 1992) operationalisiert, wobei der jeweils höhere sozioökonomische Status (HISEI) der Elternteile herangezogen wurde. Zudem wird der Bildungsabschluss der Eltern – gemessen über den „International Standard Classification of Education“ (ISCED-97; Unesco 1997) – herangezogen. Auch hier wird der jeweils höhere Bildungsabschluss der Elternteile (HISCED) berücksichtigt. Die ursprüngliche Unterteilung in sechs Kategorien wurde in dieser Arbeit dichotomisiert. Die zwei Kategorien differenzieren Elternteile mit Berufsausbildung und/oder Studium von Eltern-teilen ohne Berufsausbildung und/oder Studium.

Das normative Lernumfeld wird in den deskriptiven Analysen über die Klassenmittelwerte verschiedener affektiver Schülermerkmale (Einstellungen, Arbeitsmoral, Motivation und Interesse) beschrieben. Zudem werden die Einschätzungen zur Schüler-Lehrerbeziehung, ebenfalls als Klassenmittelwerte, herangezogen. Alle Konstrukte zur Beschreibung des kognitiven Lernumfelds wurden mit Likert-Items gemessen.

6.2.2 Skalen der Lehrenden

Unterrichtsprozesse werden in PISA 2012 auch aus Sicht der Mathematiklehrkräfte erhoben. Konkret machen die Lehrenden Angaben zum Potential zur kognitiven Aktivierung, welches mit drei Subdimensionen erhoben wurde. Zudem liegen Daten zu den Überzeugungen der Mathematiklehrkräfte zu den Zielen des Mathematikunterrichts vor. Hier werden vier Zielbereiche unterschieden (siehe Tabelle 5). Beide Skalen wurden im Rahmen der COACTIV-Studie (Kunter, Baumert et al. 2011) entwickelt. Alle Dimensionen, welche das Potential zur kognitiven Aktivierung und zu den Zielen des Mathematikunterrichts beschreiben, wurden mit Likert-Items gemessen.

Tabelle 5. Skalen der Lehrenden aus PISA 2012 – nationale Erweiterung (Mang et al. 2018).

Themenkomplex	Konstruktname	Bezeichnung	Likert-Format
Überzeugungen zu den Zielen des Mathematikunterrichts	Unterrichtsziel: Anwendung im Alltag	DEU_TE34_ANW	✓
	Unterrichtsziel: Modellierungsfähigkeit	DEU_TE34_MOD	✓
	Unterrichtsziel: Beherrschung von Routinen	DEU_TE34_ROU	✓
	Unterrichtsziel: Interesse	DEU_TE34_INT	✓
Potential zur kognitiven Aktivierung	Insistieren auf Erklären und Begründen	DEU_TE35_ERK	✓
	Verständnisfördernde Variation von Aufgabenstellungen	DEU_TE35_VAR	✓
	Kognitiv aktivierende Aufgaben	DEU_TE35_KOG	✓

Ein Häkchen (✓) in der letzten Spalte gibt an, dass das Konstrukt mit Likert-Items gemessen wird.

6.3 Analyseebene und Gewichtung

Die Analyseebene richtet sich nach der jeweiligen Fragestellung, weshalb hier gesondert auf die verschiedenen Analyseebenen und die jeweils verwendeten Gewichtungsvariablen eingegangen wird. Die Gewichte berücksichtigen die Stichprobenausfälle im Längsschnitt und wurden vom Zentrum für internationale Vergleichsstudien (ZIB) zur Verfügung gestellt.

Fragestellung 1 hinterfragt die Vergleichbarkeit der vorliegenden Fragebogendaten zwischen den Schularten. Daher wurde im ersten Schritt die Invarianz der Fragebogenskalen zwischen den Schularten auf Individualebene geprüft (Abschnitt 7.1.1). Hierfür wurde bei den Lernenden ein Individualgewicht einbezogen, welches die Ausfälle im Längsschnitt berücksichtigt (PPTOTWGT). Für die Lehrenden liegen keine entsprechenden Gewichte vor. Da die Anzahl der Lehrenden zwischen den Schularten jedoch stark schwankt, wurde bei den Invarianzanalysen der Lehrenden zumindest proportional gewichtet. Zusätzlich zu den Invarianzanalysen wurde im zweiten Schritt zur Beantwortung von Fragestellung 1 geprüft, inwieweit die Skalen, welche auf Likert-Items basie-

ren, durch Antworttendenzen verzerrt sind (Abschnitt 7.1.2). Für die Analyse von Antworttendenzen wurden auf Individualebene Indikatoren für Akquieszenz (ARS) und Midpoint-Responding (MPR) gebildet (siehe Abschnitt 6.4.3). Diese werden für die weiteren Analysen von Antworttendenzen auf Klassenebene verwendet. Hierfür wurden die Individualwerte als Klassenmittelwerte aggregiert, wobei ein Gewicht angewandt wurde, welches die Ausfallwahrscheinlichkeit innerhalb der Klasse berücksichtigt (PPWCSWGT). Die resultierenden Klassenmittelwerte wurden anschließend entsprechend der Ausfallwahrscheinlichkeit der Klasse und der Schule bei den deskriptiven Analysen auf Klassenebene gewichtet (PPCLSWGT).

Für Fragestellung 2 wurden die Schularten deskriptiv auf Klassenebene beschrieben (Abschnitt 7.2). Hierfür wurden die Daten der Lernenden zunächst als Klassenmittelwerte aggregiert, wobei entsprechend der Ausfallwahrscheinlichkeit innerhalb der Klasse gewichtet wurde (PPWCSWGT). Die resultierenden Mittelwerte wurden anschließend entsprechend der Ausfallwahrscheinlichkeit der Klasse und der Schule analysiert (PPCLSWGT). Für die Lehrenden liegt nur ein Wert pro Klasse bzw. Kurs vor, weshalb hier nicht aggregiert wurde. Analog zu den Schülerdaten wurde in den deskriptiven Analysen gewichtet (PPCLSWGT).

Fragestellung 3 zielt ebenfalls auf die Klassenebene. In Mediationsmodellen wurden die klassengemittelten Merkmale untersucht (Abschnitt 7.3). Analog zu den deskriptiven Analysen wurden die Daten der Lernenden zunächst als Klassenmittelwerte aggregiert, wobei für die Ausfallwahrscheinlichkeit innerhalb der Klasse gewichtet wurde (PPWCSWGT). Die resultierenden Mittelwerte wurden anschließend entsprechend der Ausfallwahrscheinlichkeit der Klasse und der Schule auf Klassenebene gewichtet (PPCLSWGT). Für die Lehrenden liegt nur ein Wert pro Klasse vor, weshalb hier nicht aggregiert wurde. Analog zu den Schülerdaten wurde in den Mediationsanalysen auf Klassenebene gewichtet (PPCLSWGT).

Die interessierenden Merkmale in Fragestellung 4 liegen ebenfalls auf Klassenebene. Zu deren Beantwortung wurden Interaktionseffekte der Schulart mit Unterrichtsprozessmerkmalen im Zusammenhang mit der Leistungsentwicklung getestet (Abschnitt 7.4). Da die Leistung zum zweiten Messzeitpunkt jedoch unter Kontrolle der Leistung zum ersten Testzeitpunkt analysiert wurde, sind Mehrebenen-Modelle angemessen. Die genauen Modelle werden im Me-

thodenteil in Abschnitt 6.9 näher spezifiziert. Gewichtet wurde auf Individual-ebene (PPWCSWGT) und auf Klassenebene (PPCLSWGT). Auf Klassenebene wurden wiederum die Klassenmittelwerte analysiert. Analog zu den deskriptiven Analysen wurden die Daten der Lernenden zunächst als Klassenmittelwerte aggregiert, wobei für die Ausfallwahrscheinlichkeit innerhalb der Klasse gewichtet wurde (PPWCSWGT). Die resultierenden Mittelwerte wurden anschließend entsprechend der Ausfallwahrscheinlichkeit der Klasse und der Schule auf Klassenebene gewichtet (PPCLSWGT). Für die Lehrenden liegt nur ein Wert pro Klasse vor, weshalb hier nicht aggregiert wurde.

6.4 Vergleichbarkeit der Fragebogeninstrumente

Für den Vergleich von Merkmalen und Modellen zwischen den Schularten gilt es vorab zu prüfen, ob die eingesetzten Messinstrumente äquivalent sind (siehe Kapitel 4). Hierfür wurde für die Lernenden mittels mehrgruppenkonfirmatorischer Faktorenanalyse (MGCFA) geprüft, ob die Zusammenhänge der Variablen (Faktoren) in den Vergleichsgruppen invariant sind.

Für die Daten der Lehrenden werden hingegen die Faktorstrukturen aus Hauptkomponentenanalysen (PCA) analysiert und zwischen den Schularten verglichen. Invarianzanalysen mittels MGCFA können für die Daten der Lehrenden nicht realisiert werden, da die Fallzahlen zu gering sind. Die geschätzten Parameter sowie die Standardfehler wären nicht belastbar (Wolf et al. 2013). Die PCA eignet sich hingegen auch bei kleiner Fallzahl, da hier keine Parameter geschätzt, sondern die Faktoren direkt aus den erhobenen Daten berechnet werden.

6.4.1 Invarianzanalyse mittels mehrgruppenkonfirmatorischer Faktorenanalyse

Die mehrgruppenkonfirmatorische Faktorenanalyse (MGCFA) ist eine Erweiterung der konfirmatorischen Faktorenanalyse (CFA) (Brown 2006). In der einfachen CFA werden theoretische Annahmen über die Zusammenhänge zwischen beobachteten Variablen und latenten Variablen geprüft. Eine latente Variable ist ein Konstrukt, das nicht direkt gemessen werden kann (bspw. die

wahrgenommene Unterstützung durch den Lehrenden). Es lassen sich jedoch Rückschlüsse aus erfassbaren Indikatoren (bspw. Fragebogenitems) ziehen (Backhaus et al. 2016). Ist die wahrgenommene Unterstützung durch den Lehrenden hoch, ist anzunehmen, dass ein Lernender die Frage nach dem Interesse des Lehrenden für den Lernfortschritt positiv beantwortet. Die Annahmen über diese Zusammenhänge zwischen latenter Variable (η) und den erfassbaren Indikatoren (X) zielen also auf das Ladungsmuster der Variablen (Λ). Zudem können Annahmen über die Metrik der Ladungen, gemessen über Intercepts (τ_x), gemacht werden. Varianz der Indikatoren, welche nicht über die latente Variable (η) erklärt werden kann, wird als Residualvarianz der Indikatoren (ε) in das Modell aufgenommen (Brown 2006; Steenkamp und Baumgartner 1998).

$$X = \tau_x + \Lambda_x \eta + \varepsilon \quad (1)$$

Mit der CFA wird überprüft, ob die empirisch ermittelten Zusammenhänge von den hypothetischen Annahmen abweichen. Hierfür werden verschiedene Modellparameter so geschätzt, dass die zugehörige geschätzte Kovarianzmatrix größtmögliche Übereinstimmung mit der Kovarianzmatrix der empirischen Daten hat (Brown 2006).

Die mehrgruppenkonfirmatorische Faktorenanalyse erlaubt es, die Modellparameter parallel in verschiedenen Gruppen (hier den Schularten) zu schätzen (für eine genaue Beschreibung des Mehrgruppenmodelles, des Schätzverfahrens sowie der Modellannahmen siehe Millsap 2011). Damit kann geprüft werden, wie sehr die geschätzten Parameter (bspw. die Faktorladungen) zwischen den Schularten übereinstimmen bzw. invariant sind (Cheung und Rensvold 1999).

Die Invarianzanalyse erfolgt in einer hierarchischen Abfolge. Schrittweise werden die geschätzten Parameter zwischen den Gruppen gleichgesetzt. Hierdurch werden die drei Stufen der Invarianz getestet. Zunächst wird die Dimensionalität der Faktoren geprüft (*konfigurale Invarianz*), d.h., die Variablen müssen in allen Gruppen auf dem jeweils theoretisch angenommenen Faktor laden (Cheung und Rensvold 1999). Es wird also geprüft, ob die Items in allen Gruppen dieselbe Konfiguration aufweisen. Dieses Level der Invarianz gibt an, dass ähnliche, aber nicht zwingend identische Konstrukte in den Gruppen gemessen wurden (Chen 2007). Untenstehende Grafik gibt jeweils ein Beispiel für konfi-

gurale Invarianz und konfigurale Varianz bei zwei Gruppen in Anlehnung an Cheung und Rensvold (2000) wieder.

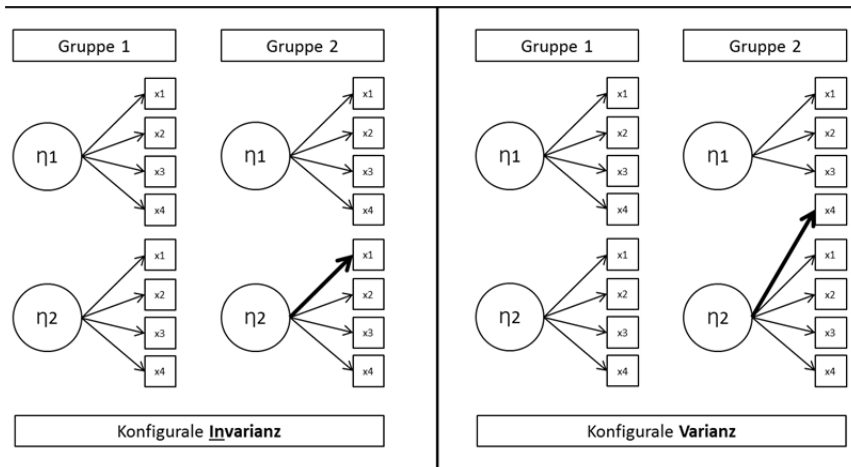


Abbildung 6. Gegenüberstellung konfiguraler Invarianz und konfiguraler Varianz.

In Anlehnung an Cheung und Rensvold (2000).

Liegt konfigurale Invarianz vor, kann die nächste Stufe der Invarianz getestet werden (Davidov 2008). Bei der *metrischen Invarianz* wird die Höhe der Ladung zusätzlich berücksichtigt. Hierfür werden die Ladungsparameter gleichgesetzt ($\Lambda_{GY} = \Lambda_{IGS} = \Lambda_{RS} = \Lambda_{MBG}$; Cheung und Rensvold 1999, 2000; Davidov 2008). Liegt metrische Invarianz vor, hat ein latenter Faktor die gleiche Maßeinheit über die Gruppen hinweg (Chen 2007). Anschließend wird im dritten Schritt auf *skalare Invarianz* getestet, indem zusätzlich die Intercepts gleichgesetzt werden ($\tau_{GY} = \tau_{IGS} = \tau_{RS} = \tau_{MBG}$; Cheung und Rensvold 2000). Dies impliziert, dass Unterschiede in den Itemmittelwerten tatsächlich durch Unterschiede in den Mittelwerten der latenten Variable hervorgerufen werden (Davidov 2008).

Die drei Stufen der Invarianz werden für jede Skala mittels Mplus 7 (Muthén und Muthén 1998–2013) unter Berücksichtigung der Mehrebenenstruktur geschätzt (Type is complex). Zur Modellidentifikation muss dem latenten Faktor eine Metrik zugewiesen werden. Hierfür wird jeweils die erste Ladung eines Faktors auf den Wert eins gesetzt (siehe hierzu auch Steenkamp und Baum-

gartner 1998). Das Messniveau der Items wird als kontinuierlich betrachtet. Dies wird teils kritisch diskutiert (siehe bspw. Lubke und Muthén 2004). Simulationsstudien zeigen jedoch, dass Mehrgruppenanalysen mit Likert-Items und nicht normalverteilten Daten die Wahrscheinlichkeit einer falschen Schlussfolgerung nicht erhöhen (Beuckelaer 2005; Davidov 2008).

Die Beurteilung der Modellpassung kann klassisch über den Chi-Quadrat-Test erfolgen. Da dieser jedoch sensibel bezüglich der Fallzahl sowie gegenüber Verletzungen der Normalverteilungsannahmen ist, empfiehlt es sich, Indizes heranzuziehen, welche zwar auf dem Chi-Quadrat-Test aufbauen, aber die genannten Nachteile überwinden können (Chen 2007; Rutkowski und Svetina 2013): Der RMSEA (Root Mean Square Error of Approximation) sowie der SRMR (Standardized Root Mean Square Residual) zählen zu den „Absolut fit indices“. Diese messen die Abweichung zwischen der empirischen Kovarianzmatrix und der geschätzten Kovarianzmatrix. Je kleiner der Wert dieser Indizes ist, desto besser ist die Modellpassung. Die RMSEA berücksichtigt zudem die Sparsamkeit des Modells, d.h. Modelle mit weniger frei geschätzten Parametern (höhere Stufe der Invarianz) bekommen bei gleicher Abweichung zwischen empirischer und geschätzter Kovarianzmatrix einen geringeren Wert und haben somit eine bessere Modellpassung. Ein Wert von 0 gibt eine optimale Passung an. Zudem prüfen „incremental fit indices“, wie der CFI (Comparative Fit Index), inwieweit das getestete Modell einem alternativen Modell überlegen ist, welches Unabhängigkeit der Indikatoren annimmt. Der Wertebereich des CFI liegt zwischen 0 und 1, wobei der Wert 1 eine perfekte Modellpassung angibt (Brown 2006). Zusätzlich kann das Bayesian-Information-Criterion (BIC) hinzugezogen werden. Dieses basiert nicht auf dem Chi-Quadrat-Test, sondern auf der Log-Likelihood und bestraft unnötig geschätzte Parameter stark. Wird nach diesem Kriterium ein Modell mit mehr Restriktionen (und damit weniger frei geschätzten Parametern und höherer Stufe der Invarianz) abgelehnt, ist es sehr sicher, dass das Modell mit weniger Restriktionen (mehr frei geschätzten Parametern) besser passt. Je niedriger der Wert des BIC, desto besser passt das Modell.

Die konfigurale Invarianz kann direkt über diese Fit Indices beurteilt werden. Für den CFI sollten die Werte über 0,95 liegen. Für den RMSEA gilt ein Wert bis maximal 0,10, für den SRMR gilt ein Wert bis 0,08 als gerade noch akzeptabel (Rutkowski und Svetina 2013). Der BIC kann für die Beurteilung konfi-

guraler Invarianz nicht einbezogen werden. Er fungiert nur als Maß zum Modellvergleich.

Metrische und skalare Invarianz kann angenommen werden, wenn das Modell mit mehr Restriktionen deutlich besser passt als das Modell mit weniger Restriktionen (Davidov 2008). Hier werden also genestete Modelle getestet und die Veränderung in den Fit Indices einbezogen. Metrische Invarianz ist gegeben, wenn der Wert der RMSEA maximal um 0,03 ansteigt bzw. der Wert der SRMR maximal um 0,01 ansteigt. Für den CFI gilt eine Verminderung von bis zu 0,02 als akzeptabel (Rutkowski und Svetina 2013). Für die Beurteilung genesteter Modelle kann zudem der BIC herangezogen werden. Ein Anstieg des BIC um maximal 10 Einheiten gilt in Anlehnung an Kaas und Raftery (1995) hier noch als akzeptabel, um metrische Invarianz anzunehmen. Für skalare Invarianz empfiehlt es sich, teils strengere Grenzwerte zu wählen. Hier gilt ein Anstieg um 0,01 für RMSEA und SRMR als akzeptabel, sowie eine Reduktion von 0,01 für den CFI (Chen 2007; Rutkowski und Svetina 2013). Für den BIC gelten die gleichen Grenzen wie bei der metrischen Invarianz von maximal 10 Einheiten.

6.4.2 Faktorstrukturvergleich

Das primäre Ziel der Hauptkomponentenanalyse (Principal Component Analysis; PCA) besteht darin, einem größeren Variablensatz eine ordnende Struktur zu geben. Dabei wird angenommen, dass sich die Varianz der Variablen komplett durch die Faktoren erklären lässt. Die PCA berechnet daher die Faktoren als direkte Linearkombinationen der Indikatoren basierend auf den Itemkorrelationen (Matsunaga und Masaki 2010; Noack 2007). Durch Rotation der ursprünglichen Faktorenlösung wird anschließend die Interpretierbarkeit der Faktoren erhöht: Hohe Ladungen der Variablen auf die Faktoren werden verstärkt, niedrige Ladungen weiter abgeschwächt (Abdi 2004).¹⁴ Für die Rotation werden in dieser Arbeit korrelierte Faktoren angenommen, da die einzelnen Dimensionen der Ziele im Mathematikunterricht sowie die Dimensionen der kognitiven Aktivierung inhaltliche Überschneidungen aufweisen. Alltagsbezogene Aufgaben können bspw. auch die Fähigkeit zum Mathematisieren erfor-

14 Einen gute Einführung zur Rotation und den verschiedenen Verfahren gibt Noack (2007).

dern. Trotzdem handelt es sich hier um unterschiedliche Ziele des Mathematikunterrichts (Mathematik im Alltag anwenden vs. komplexes Modellieren üben). Um die Annahme korrelierter Dimensionen zu berücksichtigen, werden die Anfangslösung über die Funktion Oblimax rotiert, welche sich auch für kleine Fallzahlen eignet. Die Analysen werden mit dem Programm SPSS 22 (IBM Cooperation 2013) durchgeführt.

Über den Scree Plot (Cattell 1966) sowie das Kaiser-Gutmann-Kriterium (Guttman 1954) wird über die Anzahl der Dimensionen entschieden. Beide Kriterien beziehen sich auf die Eigenwerte der Faktoren. Der Eigenwert eines Faktors gibt an, wieviel von der Gesamtvarianz aller Variablen durch diesen Faktor erklärt wird. Ist ein Eigenwert kleiner als 1, erklärt der Faktor weniger als die Varianz einer einzigen Variablen. Nach dem Kaiser-Gutmann-Kriterium werden alle Faktoren, deren Eigenwert kleiner 1 ist, als unbedeutend erklärt. Der Scree Plot visualisiert die Eigenwerte der Faktoren, welche in absteigender Reihenfolge abgebildet werden. Mit zunehmender Anzahl an Faktoren sinkt die Varianzaufklärung eines zusätzlichen Faktors, was sich durch eine abflachende Kurve zeigt. Durch optische Begutachtung wird ermittelt, an welchem Punkt die Kurve abflacht. Faktoren, die links dieses sogenannten „Knicks“ liegen werden als relevant betrachtet. Das heißt, auch Faktoren, deren Eigenwert größer als 1 ist, können über die subjektive Betrachtung des Scree Plots aus der weiteren Analyse ausgeschlossen werden (Bortz und Schuster 2010; Costello und Osborne 2005).

Um die Vergleichbarkeit zwischen den Schularten zu prüfen, werden PCAs – für die Ziele des Mathematikunterrichts sowie für die kognitive Aktivierung – über alle Lehrenden hinweg berechnet und zusätzlich innerhalb der Schularten. Über die Komponentenmatrizen wird dann das Ladungsmuster der Items auf den Faktoren zwischen den Gruppen verglichen. Laden Items nicht eindeutig auf einem Faktor oder zeigen sich Unterschiede in den Ladungsmustern zwischen den Gruppen, werden die entsprechenden Items aus der Analyse ausgeschlossen (siehe hierzu auch Matsunaga und Masaki 2010). Mit den verbleibenden Items wird die Analyse wiederholt, so lange, bis eindeutige Ladungen sowie vergleichbare Ladungsmuster zwischen den Gruppen gegeben sind.

Zusätzlich wird der Kontingenzkoeffizient nach Tucker (1951) berechnet. Dieser gibt einen objektiven Wert für die Ähnlichkeit zweier Faktoren an. Hierfür

werden jeweils die Ladungen des Modells innerhalb einer Schulart (λ_{SAi}) mit den Ladungen der Gesamtstichprobe (λ_{Gi}) pro Faktor in Bezug gesetzt.

$$C_{jk} = \frac{\sum_{i=1}^p \lambda_{SAi} \lambda_{Gi}}{\sqrt{(\sum_{i=1}^p \lambda_{SAi}^2)(\sum_{i=1}^p \lambda_{Gi}^2)}} \quad (2)$$

Je höher beide Faktorladungen λ_{SAi} und λ_{Gi} sind, desto größer ist der Zähler. Bei vergleichbaren Faktormustern sollte bei einer hohen Ladung innerhalb einer Schulart auch eine hohe Ladung in der Gesamtstichprobe bei der gleichen Variablen auftreten. Der Zähler wird negativ, wenn ein Großteil der Faktorladungen gegenpolig ist, bspw. wenn in der Gesamtstichprobe positive Ladungen auftreten und innerhalb der Schulart negative Ladungen. Im Nenner wird die Varianzaufklärung des Faktors innerhalb einer Schulart mit der Varianzaufklärung des Faktors in der Gesamtstichprobe multipliziert. Je höher die Varianzaufklärung der Faktoren ist, desto stärker müssen die Übereinstimmungen der Ladungen zwischen den Vergleichsfaktoren sein. Der Kontingenzkoeffizient kann somit einen Wertebereich zwischen -1 und 1 annehmen. Ein stark negativer Wert tritt auf, wenn beide Faktoren eine hohe Varianzaufklärung aufweisen, die Faktorladungen jedoch gegenpolig sind. Hohe positive Werte treten hingegen auf bei hoher Varianzaufklärung und vergleichbaren Faktorladungen. Ab einem Wert von $C_{jk} > 0,80$ kann von vergleichbaren Faktorstrukturen ausgegangen werden (Barrett 1986).

Einige Autoren empfehlen, die Faktorstrukturen der Gruppenlösungen auf eine Zielstruktur hin zu rotieren. Dies findet in der Praxis jedoch selten Anwendung. Ohne Rotation kann die Vergleichbarkeit der Faktorstrukturen unterschätzt werden. Bei hohen Kongruenzkoeffizienten führt eine Rotation hin zur Zielstruktur jedoch zu keinen nennenswerten Unterschieden (van de Vijver und Leung 1998). Die Koeffizienten liegen für die Ziele im Mathematikunterricht sowie für die kognitive Aktivierung deutlich über dem Grenzwert von $0,80$, weshalb hier auf die Faktorrotation verzichtet wurde.

Entsprechend der finalen Faktorenlösung wird für die Lehrenden pro Faktor ein Mittelwert aus den zugehörigen Items gebildet. Damit erhält jeder Lehrende für jeden Faktor einen Wert, welcher die Relevanz der Ziele im Mathematikunterricht bzw. die Praktiken zur kognitiven Aktivierung im Unterricht für jede Dimension wiedergibt. Die resultierenden Variablen sind Bestandteil der weite-

ren Analysen. Das verwendete Analyseprogramm SPSS 22 (IBM Cooperation 2013) bietet auch die Möglichkeit basierend auf der PCA Faktorwerte auszugeben. Hierbei werden jedoch unkorrelierte Faktoren erzwungen, was inhaltlich hier nicht gerechtfertigt ist (Fromm 2010).

6.4.3 Berechnung von Indikatoren für Antworttendenzen

In den eingesetzten Fragebogeninstrumenten der PISA-Studie 2012 finden sich im Schülerfragebogen 13 Einstellungs- und Verhaltensfragen mit insgesamt 81 Items im Likert-Format (*1=stimme völlig zu, 2=stimme eher zu, 3=stimme eher nicht zu, 4=stimme überhaupt nicht zu*). Im Lehrerfragebogen werden 17 Einstellungs- und Verhaltensfragen mit insgesamt 164 Items im Likert-Format gestellt. Bei 13 Fragen bzw. 142 Items ist die Polung der Antwortskala von zustimmend zu ablehnend (*1=trifft zu, 2=trifft eher zu, 3=trifft eher nicht zu, 4=trifft nicht zu* bzw. *1=stimme völlig zu, 2=stimme eher zu, 3=stimme eher nicht zu, 4=stimme überhaupt nicht zu*). 4 Fragen bzw. 22 Items haben inverse Antwortskalen (*1=stimmt überhaupt nicht, 2=stimmt eher nicht, 3=stimmt eher, 4=stimmt ganz genau*).

Der Akquieszenzindex (ARS), der zugleich Disakquieszenz misst, errechnet sich über die Differenz aus der Anzahl zustimmender Antworten und der Anzahl ablehnender Antworten, geteilt durch die Anzahl der einbezogenen Items (van Herk et al. 2016). Die ersten beiden Kategorien der Antwortskala werden als Zustimmung in die Berechnung einbezogen, die letzten beiden Kategorien als Ablehnung. Es resultiert ein Index mit einem Wertebereich von -1 bis 1 , wobei positive Werte für eine Zustimmungstendenz stehen, negative Werte für die Ablehnung von Aussagen.

Der Index, welcher die Tendenz zur Mitte misst (MPR) und zugleich das Ausmaß an Extremantworten, errechnet sich über die Summe der Antworten an den Extrempolen geteilt durch die Anzahl der einbezogenen Items (van Herk et al. 2016). Der Index hat damit einen Wertebereich von 0 bis 1 , wobei ein Wert von 1 für die ausschließliche Wahl extremer Antwortkategorien steht, ein Wert von 0 verweist auf die ausschließliche Wahl mittlerer Antwortkategorien.

Um zu gewährleisten, dass die errechneten Indizes nicht mit inhaltlich begründbaren Antwortmustern vermischt werden, muss sichergestellt werden,

dass die verwendeten Items heterogen sind, d.h., die Items sollten verschiedene, möglichst unkorrelierte Konstrukte messen (Baumgartner und Steenkamp 2001). Die mittleren Interkorrelationen der Items geben hierüber Aufschluss. Zu deren Berechnung werden die Items aller Fragen miteinander korreliert und anschließend der Mittelwert der einzelnen Korrelationen berechnet. Da Korrelationswerte (r) nicht intervallskaliert sind, darf der Mittelwert der Korrelationen jedoch nicht direkt berechnet werden. Stattdessen werden die einzelnen Korrelationen in Fisher Z-Werte (Z) transformiert und der Mittelwert dieser transformierten Werte berechnet. Anschließend wird dieser wieder in eine Korrelation zurücktransformiert (Bortz und Schuster 2010; Silver und Dunlap 1987). Die Transformationen erfolgen nach folgenden Gleichungen:

$$r_{yx} = \frac{e^{2Z} - 1}{e^{2Z} + 1} \quad (3)$$

bzw.

$$Z = \frac{1}{2} * \ln\left(\frac{1 + r}{1 - r}\right) \quad (4)$$

Für die Schülerantworten ergibt sich für die beschriebenen PISA-2012-Items eine mittlere Interkorrelation von $r_{yx}=0,06$. Für die Lehrerantworten liegt die mittlere Interkorrelation bei $r_{yx}=0,03$. Als homogen gelten Items mit einer mittleren Interkorrelation von 0,20 bis 0,40 (Briggs und Cheek 1986). Im Umkehrschluss werden heterogene Items angenommen, wenn die mittlere Interkorrelation deutlich unter einem Wert von 0,20 liegt. Dies ist für die ermittelte mittlere Interkorrelation eindeutig gegeben, weshalb die berechneten Indizes für ARS und MPR nicht mit inhaltlich begründbaren Antwortmustern vermischt sind.

Die Analyse von Mittelwerten und Verteilungen der Indizes erlaubt Rückschlüsse auf die Ausprägung der Antworttendenzen der Lernenden sowie der Lehrenden. Inwieweit Schulartunterschiede in den Antworttendenzen bestehen und welche Merkmale besonders anfällig für Verzerrungen sind, kann mit der einfaktoriellen Varianzanalyse bzw. der Kovarianzanalyse geprüft werden (siehe Abschnitt 6.5).

6.5 Varianzanalytische Modelle für die Analyse von Antworttendenzen und die Deskription der Schularten

In der vorliegenden Arbeit werden für die deskriptiven Beschreibungen der Schularten alle Schülerdaten auf Klassenebene aggregiert, da der Fokus auf Prozessen des Unterrichtes sowie damit verbundenen Gruppenprozessen liegt. Die Lehrenden können ebenfalls je einer Klasse zugeordnet werden. Die Merkmale wurden zunächst auf Individualebene z-standardisiert, um sie um den Mittelwert aller befragten Lernenden bzw. der Mathematiklehrenden zu zentrieren. Die Daten der Lernenden und Lehrenden werden als Mittelwert der Klassen aggregiert. In diesem Schritt wird auch entsprechend der Ausfallwahrscheinlichkeit der Lernenden gewichtet, d.h., Daten von Lernenden, welche eine höhere Ausfallwahrscheinlichkeit haben, bekommen bei der Berechnung des Klassenmittelwertes ein höheres Gewicht (siehe Abschnitt 6.3). Diese aggregierten Daten werden dann pro Schulart gemittelt, wobei hier für die Ziehungs- und Ausfallwahrscheinlichkeit der Klasse gewichtet wird. Auch für die Lehrenden werden die Daten entsprechend der Ausfallwahrscheinlichkeit der Klasse gemittelt. Diese Schulartmittelwerte sind dann Ausgangspunkt für die Beschreibung der Schularten.

Ein direkter Vergleich dieser Mittelwerte zwischen den Schularten führt jedoch zu keinem befriedigenden Urteil, da es keine Anhaltspunkte gibt, welche Abweichungen zwischen den Gruppen noch akzeptabel sind, um gleiche oder ungleiche Mittelwerte anzunehmen (Rasch et al. 2008). Stattdessen bieten sich varianzanalytische Verfahren an, um die Bedeutung von Mittelwertunterschieden zu beurteilen. In der vorliegenden Arbeit wird jeweils die Varianz eines Merkmals (abhängige Variable) zwischen den Schularten (Gruppierungsvariable) betrachtet. Gruppierungsvariablen werden in der Varianzanalyse auch Faktoren genannt, weshalb das hier eingesetzte Verfahren mit einer Gruppierungsvariablen bzw. einem Faktor als einfaktorielle Varianzanalyse bezeichnet wird (Handl 2010; Rasch et al. 2008). Die Kovarianzanalyse ist eine Erweiterung dieses Verfahrens. Sie wird verwendet, wenn eine weitere metrische Variable einen Einfluss auf das untersuchte Merkmal hat und dieser Einfluss herausgerechnet werden soll (Bortz und Schuster 2010). Antworttendenzen können beispielsweise die Ausprägungen der untersuchten Merkmale ungewollt beeinflussen. Dieser Einfluss muss herausgerechnet werden, um Mittelwerte zwischen den Schularten vergleichen zu können. Im Folgenden werden zunächst die Grundlagen der einfaktoriellen Varianzanalyse beschrieben sowie das Ef-

fektmäß Eta Quadrat (η^2) zur Beurteilung von Mittelwertunterschieden. Anschließend wird die Erweiterung zur Kovarianzanalyse vorgestellt.

6.6 Einfaktorielle Varianzanalyse

Die einfaktorielle Varianzanalyse untersucht, in welchem Ausmaß die Variabilität eines Merkmals zur Beschreibung der Schularten (bspw. die Antworten von Befragten auf ein Item oder die Tendenz, Fragen zuzustimmen,) eher auf die Schulart zurückzuführen ist, oder ob die Variabilität durch andere Einflüsse bedingt ist. Ist der Anteil an Variabilität, die auf den untersuchten Faktor zurückzuführen ist, genügend groß, wird angenommen, dass sich auch die zugehörigen Mittelwerte bedeutend unterscheiden (Diehl 1983).

Die Gesamtvariabilität des Merkmals (QS_{Ges}) wird ermittelt, indem für jeden Messwert (y_i) die quadrierte Abweichung zum Mittelwert (\bar{y}) berechnet wird und diese Abweichungen aufsummiert werden.¹⁵

$$QS_{Ges} = \sum_{i=1}^n (y_i - \bar{y})^2 \quad (5)$$

Nun kann der Anteil der Variabilität ermittelt werden, die auf den Faktor (die Schulart) zurückzuführen ist. Rechnerisch wird die Gesamtvariabilität (QS_{Ges}) zerlegt in die Variabilität innerhalb der Gruppen ($QS_{innerhalb}$) und die Variabilität zwischen den Gruppen ($QS_{zwischen}$).

$$QS_{Ges} = QS_{innerhalb} + QS_{zwischen} \quad (6)$$

Die *Variabilität innerhalb der Gruppen* wird auch als Residualvariabilität bezeichnet. Sie berechnet sich als die quadrierten Abweichungen der individuellen Werte (y_i) vom jeweiligen Gruppenmittelwert (\bar{A}).

15 Die Beschreibung der einfaktoriellen Varianzanalyse erfolgt in Anlehnung an Rasch et al. (2008).

$$\begin{aligned}
 QS_{\text{innerhalb}} &= (y_{iGY} - \bar{A}_{GY})^2 + (y_{iIGS} - \bar{A}_{IGS})^2 \\
 &+ (y_{iRS} - \bar{A}_{RS})^2 \\
 &+ (y_{iMBG} - \bar{A}_{MBG})^2
 \end{aligned} \tag{7}$$

Unter der Annahme, dass die Variabilität der Merkmale gänzlich auf die Schulart zurückzuführen ist, dürften sich die Merkmalsausprägungen (wie die Antworten der Befragten auf ein Item oder die Tendenz, Fragen zuzustimmen) innerhalb der Gruppen nicht unterscheiden ($QS_{\text{innerhalb}} = 0$). Finden sich hingegen Unterschiede innerhalb der Gruppen, gibt es über die Schulart hinaus Einflussgrößen, welche die Ausprägungen der Merkmale beeinflussen ($QS_{\text{innerhalb}} > 0$).

Die *Variabilität zwischen den Gruppen* bestimmt den Anteil der Unterschiedlichkeit aller Messwerte, der auf die Schularten zurückzuführen ist. Hierfür werden die quadrierten Abweichungen der jeweiligen Gruppenmittelwerte (\bar{A}) vom Gesamtmittelwert (\bar{G}) des Merkmals aufsummiert.

$$\begin{aligned}
 QS_{\text{zwischen}} &= (\bar{A}_{GY} - \bar{G})^2 + (\bar{A}_{IGS} - \bar{G})^2 \\
 &+ (\bar{A}_{RS} - \bar{G})^2 + (\bar{A}_{MBG} - \bar{G})^2
 \end{aligned} \tag{8}$$

Basierend auf den beschriebenen Variabilitätskomponenten ($QS_{\text{innerhalb}}$ und QS_{zwischen}) lassen sich nun verschiedene Maße berechnen, um den Anteil der Schulart an der Variabilität der Merkmalsausprägungen (bspw. der Antworten von Befragten) quantifizieren zu können. In dieser Arbeit wird das Effektmaß Eta Quadrat (η^2) verwendet, welches sich als zuverlässig bei großen Stichproben erwiesen hat und sich direkt mit der Statistiksoftware SPSS 22 (IBM Cooperation 2013) berechnen lässt (Bortz und Schuster 2010). Eta Quadrat setzt die Variabilität zwischen den Gruppen ins Verhältnis zur Variabilität innerhalb der Gruppen addiert mit der Variabilität zwischen den Gruppen.

$$\eta^2 = \frac{QS_{\text{zwischen}}}{QS_{\text{zwischen}} + QS_{\text{innerhalb}}} \tag{9}$$

Gäbe es keinen Effekt auf die Merkmalsausprägungen über die Schulart hinaus, hätte Eta Quadrat einen Wert von 1. Je größer jedoch die Variabilität innerhalb der Gruppen ist (also andere Größen die Gesamtvariabilität beeinflussen), desto größer ist die Residualvarianz $QS_{\text{innerhalb}}$ und desto geringer ist der Effekt der

Schulart auf die Merkmalsausprägungen. Dies spiegelt sich in entsprechend kleineren Werten für Eta Quadrat wider.

Zur Interpretation von η^2 gelten folgende Grenzen (Bortz & Schuster 2010): 0,01 bis 0,05 zeigt einen kleinen Effekt, 0,06 bis 0,14 steht für einen mittleren Effekt und Werte über 0,14 bezeichnen einen starken Effekt. Die Effektstärken können auch als prozentualer Anteil der Schulart an der Varianz eines Merkmals interpretiert werden. Ein η^2 von 0,14 bedeutete beispielsweise, dass 14 % der Varianz eines Merkmals durch die Schulart erklärt werden.

6.7 Kovarianzanalyse

Die Kovarianzanalyse wird als Erweiterung der einfaktoriellen Varianzanalyse verwendet, wenn zusätzlich für einen metrischen Faktor (die Kovariate) kontrolliert werden soll. Damit erlaubt die Kovarianzanalyse bspw. den Einfluss von Störvariablen herauszurechnen sowie den Effekt der Kovariablen zu bestimmen (Bortz und Schuster 2010). In dieser Arbeit werden die Antworttendenzen der Befragten als Kovariate in das Modell aufgenommen. Zeigen sich starke Effekte der Kovariaten, werden in den weiteren Analysen die Mittelwerte unter Kontrolle der Antworttendenzen berichtet.

Rechnerisch stellt die Kovarianzanalyse eine Kombination aus Regressionsanalyse und Varianzanalyse dar. Zunächst wird die gemeinsame Variation der abhängigen Variablen mit der Kovariaten berechnet ($QS_{\hat{y}}$). Dadurch wird der Anteil an Variation in der abhängigen Variablen ermittelt, welcher durch die Kovariate erklärt werden kann (Rasch et al. 2008). Die zugehörige Regressionsgleichung für die Zustimmungstendenz lautet:

$$\hat{y}_i = b_0 ARS_i \quad (10)$$

Über diese Gleichung lassen sich für jede Person i Werte für \hat{y} bestimmen (für die exakte Berechnung siehe Moosbrugger 1978). Die gemeinsame Variation errechnet sich in Anlehnung an (5) wie folgt¹⁶:

16 Die Beschreibung der zweifaktoriellen Varianzanalyse erfolgt in Anlehnung an Bortz und Schuster (2010).

$$QS = \hat{y}_i - \frac{\sum \hat{y}_i}{n} \quad (11)$$

Der gemeinsamen Variation gegenüber steht Variation in der abhängigen Variablen, welche nicht durch die Kovariate erklärt wird, die Variation der Residuen (s_e). Die Residuen berechnen sich als Abweichung der Personenwerte y_i vom geschätzten Wert (\hat{y}_i)

$$e_i = \sqrt{(y_i - \hat{y}_i)^2} \quad (12)$$

Die Variation der Residuen errechnet sich in Anlehnung an (5) wie folgt:

$$QS_e = \sum (e_i - \bar{e})^2 \quad (13)$$

Genau dieser Teil der Variation, der nicht durch die Kovariate (ARS) erklärt wird, ist dann Basis für die Varianzanalyse mit der abhängiger Variablen und dem Faktor (hier der Schulart) in Anlehnung an Abschnitt 6.6. Zur Vereinfachung wurde in der Beschreibung nur eine Antworttendenz aufgenommen. In den Analysen wurde hingegen für zwei Antworttendenzen (ARS und MPR) gleichzeitig kontrolliert.

6.8 Mediationsanalysen: Vermittelnde Mechanismen für Schulartunterschiede in den Unterrichtsprozessen

Mit Mediationsanalysen wird untersucht, ob ein direkter Effekt zwischen zwei Variablen (X und Y) durch den Einfluss einer weiteren Variablen (der sog. „Mediatorvariablen“) unterbrochen wird. Eine Mediatorvariable (M) vermittelt also zwischen zwei Variablen und spezifiziert den Mechanismus, der die Einflussbeziehung zwischen X- und Y-Variable erzeugt (Urban und Mayerl 2018). Zur Beantwortung von Fragestellung 3 werden daher Mediationsanalysen berechnet, um zu ergründen, ob der Effekt der Schulart auf die Unterrichtsprozessmerkmale vermittelt wird bzw. mediiert wird über Merkmale der Schülergruppe, der Lehrenden oder der Unterrichtsinhalte.

Bei der Mediationsanalyse wird unterschieden zwischen direkten und indirekten Effekten. Direkte Effekte bezeichnen direkte Einflussbeziehungen, die nicht durch dritte Variablen unterbrochen bzw. interveniert werden (Urban und Mayerl 2018). In Abbildung 7 repräsentiert der Pfad β_{i_M} den direkten Effekt der

Schulart auf den Mediator, wobei für die Schularten (i) Realschule, integrierte Gesamtschule und Schulen mit mehreren Bildungsgängen jeweils getrennte Werte geschätzt werden, welche den Unterschied in der Ausprägung des Mediators für die jeweilige Schulart im Vergleich zum Gymnasium wiedergeben. Pfad β_M repräsentiert den direkten Effekt des Mediators auf das Unterrichtsprozessmerkmal, unter Kontrolle der Schulart. Pfad (β_i) repräsentiert den direkten Effekt der Schulart auf das Unterrichtsprozessmerkmal, wobei wiederum für die Schularten (i) Realschule (RS), integrierte Gesamtschule (IGS) und Schulen mit mehreren Bildungsgängen (MBG) jeweils getrennte Werte geschätzt werden. Das Produkt aus den Pfaden β_{i_M} und β_M ($\beta_{i_M} * \beta_M = \beta_{i_IND}$) repräsentiert den mediierten bzw. indirekten Effekt der Schulart auf das Unterrichtsprozessmerkmal. Wenn der indirekte Effekt der Schulart auf das Unterrichtsprozessmerkmal signifikant ist, wird von einer vorliegenden Mediation ausgegangen, d.h., dass der Effekt der Schulart über die Mediatorvariable vermittelt wird und somit zur Erklärung von Schulartunterschieden in den Unterrichtsprozessen bedeutsam ist (MacKinnon 2012).

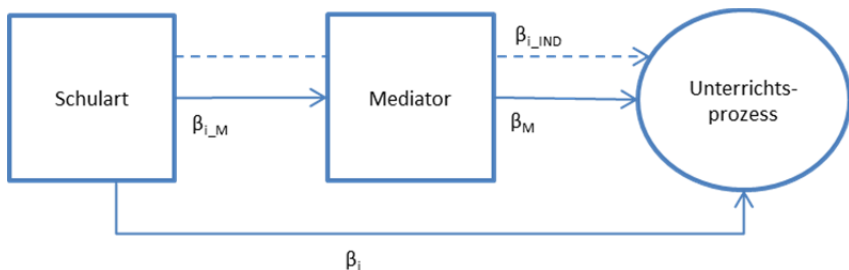


Abbildung 7. Vereinfachtes Mediationsmodell der interessierenden Effekte auf Klassenebene.

Unterschieden wird, ob eine partielle oder totale Mediation vorliegt. Eine partielle Mediation liegt vor, wenn der Effekt der Schulart auf das Unterrichtsprozessmerkmal über den Mediator führt, aber die Schulart auch einen direkten Effekt auf das Unterrichtsprozessmerkmal ausübt, der nicht durch den Mediator interveniert wird. Eine totale Mediation liegt hingegen vor, wenn der Effekt von der Schulart auf das Unterrichtsprozessmerkmal komplett durch den Mediator interveniert wird und keine direkte Beziehung mehr zwischen der Schulart und den Unterrichtsprozessmerkmal besteht (vgl. Urban und Mayerl 2018).

Berechnet werden die direkten und indirekten Effekte des beschriebenen Modells in Abbildung 7 mittels Regressionsanalysen.¹⁷ Ausgangspunkt ist die grundlegende Beziehung zwischen der Schulart (X) und dem Unterrichtsprozessmerkmal (Y):

$$Y = \beta_0 + c_i X_i + e_1 \quad (14)$$

Die relevante Information aus Gleichung (14) ist, dass c_i die Koeffizienten der Beziehung zwischen Schulart (x) und Unterrichtsprozessmerkmal (Y) beschreiben und dass e_1 die Varianz im Unterrichtsprozessmerkmal (Y) beschreibt, die nicht durch die Schulart (X) erklärt wird. Der Term β_0 bezeichnet die Regressionskonstante und ist hier nicht weiter relevant.

In diese Ausgangsgleichung wird nun die Mediatorvariable (M) aufgenommen:

$$Y = \beta_0 + \beta_i X_i + \beta_M M + e_2 \quad (15)$$

Für Gleichung (15) ist zu beachten, dass der Koeffizient für die Beziehung zwischen der Schulart (X) und dem Unterrichtsprozessmerkmal (Y) von dem Koeffizienten in Gleichung (14) abweicht, da er unter Kontrolle der Mediatorvariablen berechnet wird. Deshalb wird dieser hier als β_i bezeichnet. Der Effekt c in Gleichung (14) wird meist als totaler Effekt bezeichnet, er ist analog zur Summe des direkten Effekts der Schulart und des indirekten Effektes der Schulart auf das Unterrichtsprozessmerkmal. Aus dem Koeffizient β_i ist hingegen der indirekte Effekt, welcher über den Mediator vermittelt wird, „herausgerechnet“ und somit verbleibt der direkte Effekt der Schulart auf das Unterrichtsprozessmerkmal (Jose 2013).

Die Effekte der Schulart auf die Mediatorvariable werden analog zu Gleichung (14) berechnet:

$$M = \beta_0 + \beta_{i_M} M + e_1 \quad (16)$$

Aus den Gleichungen (15) und (16) können schließlich die Koeffizienten herangezogen werden, um den indirekten Effekt $\beta_{_IND}$, wie oben beschrieben als Produkt von β_{i_M} und β_M zu berechnen.

17 Die Beschreibung der Mediationsanalyse erfolgt in Anlehnung an Jose (2013).

Die Kenntnis des indirekten Effektes reicht jedoch für sich genommen nicht aus, um den vermittelnden Effekt der Mediatorvariablen beurteilen zu können. Baron und Kenny schlagen für die empirischen Analysen des Mediationseffektes ein mehrschrittiges Vorgehen vor. Zuerst muss ein signifikanter Effekt zwischen der Schulart und dem Unterrichtsprozessmerkmal vorliegen. Zweitens muss ein signifikanter Effekt zwischen Schulart und der Mediatorvariablen bestehen. Drittens muss die mediierende Variable signifikant mit dem Unterrichtsprozessmerkmal zusammenhängen, wenn die Schulart ebenfalls Prädiktor der abhängigen Variablen ist (Baron und Kenny 1986; Jose 2013; MacKinnon 2012).

Die signifikanten Zusammenhänge der Schulart mit den Unterrichtsprozessmerkmalen und den Mediatorvariablen (Schritte 1 und 2) wurden in den Varianzanalysen zur Deskription der Schulart bereits getestet. Somit werden in den Analysen zu Fragestellung 3 diese Schritte nicht wiederholt. Schritt 3 wird in den Analysen zu Fragestellung 3 umgesetzt. Analog zu den deskriptiven Analysen werden auch die Mediationsmodelle auf Klassenebene und unter Kontrolle von Antworttendenzen bei Likert-Skalen berechnet, wie in Abbildung 8 dargestellt.

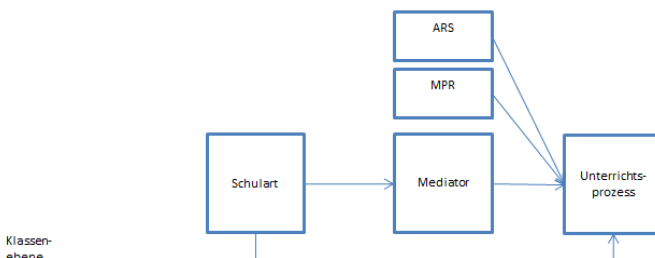


Abbildung 8. Vollständiges Mediationsmodell mit Kontrolle von Antworttendenzen.

Abkürzungen: (ARS=Akquieszenz, MPR=Midpoint Responding).

Berechnet werden die Mediationsmodelle mit der Software Mplus 7 (Muthén und Muthén 1998–2013). Die Software ermöglicht es, den indirekten Effekt mittels Bias-korrigiertem Bootstrap-Verfahren auf Signifikanz zu prüfen. Das Bootstrap-Verfahren rekonstruiert die Verteilung des interessierenden statistischen Kennwerts direkt aus einer Stichprobe, indem viele Bootstrap-Stichprobe

ben aus der untersuchten Stichprobe gezogen werden. Für jede gezogene Bootstrap-Stichprobe wird der interessierende statistische Kennwert berechnet und somit eine Verteilung der Kennwerte rekonstruiert (Shikano 2010). Das Bootstrap-Verfahren ist somit robust gegenüber Verletzungen der Normalverteilung und wird daher gegenüber anderen Verfahren, wie dem Sobel-Test, bevorzugt (Jose 2013; MacKinnon 2012). Für die Mediationsanalysen wurden in dieser Arbeit 5000 Bootstrap-Stichproben gezogen und auf dieser Basis die Konfidenzintervalle betrachtet. Das Bootstrap Konfidenzintervall gibt den Wertebereich an, in dem sich 95 % der berechneten Werte befinden. Liegt das Bootstrap-Vertrauensintervall komplett auf der positiven oder negativen Seite, enthält den Wert 0 also nicht, wird die Nullhypothese, dass kein indirekter Effekt vorliegt, verworfen (Shikano 2010).

6.9 Moderationsanalysen: Schulartunterschiede in den Zusammenhängen zwischen Unterrichtsprozessen und Leistungsentwicklung

Von einem Moderatoreffekt ist immer dann zu sprechen, wenn eine Variable (Moderatorvariable, bspw. die Schulart) den Effekt einer X-Variablen (bspw. ein Unterrichtsprozessmerkmal) auf eine abhängige Variable Y (bspw. die Leistung) beeinflusst. Eine Moderatorvariable beeinflusst also nicht die Ausprägung einer abhängigen Variablen, sondern den Effekt auf die Y-Variable. Die Moderatorvariable kann dabei die Stärke, Signifikanz und/oder Einflussrichtung des Effekts beeinflussen bzw. moderieren (Stoetzer 2017; Urban und Mayerl 2018).

Abbildung 9 zeigt auf der linken Seite das theoretische Modell (in vereinfachter Darstellung) eines Moderationseffektes der Schulart auf den Zusammenhang zwischen einem Unterrichtsprozessmerkmal und der Leistung. Für die dichotome Schulartvariable bedeutet ein Moderationseffekt, dass sich die Zusammenhänge zwischen dem Unterrichtsprozessmerkmal und der Leistung zu T2 zwischen zwei Schularten unterscheiden. Gemessen wird der moderierende Effekt der Schulart, indem zur Vorhersage der Leistung zum zweiten Messzeitpunkt neben dem Unterrichtsprozessmerkmal und der Schulart zusätzlich der Interaktionsterm als Produkt dieser beiden Merkmale aufgenommen wird. Das zugehörige Messmodell ist in Abbildung 9 auf der rechten Seite abgebildet.

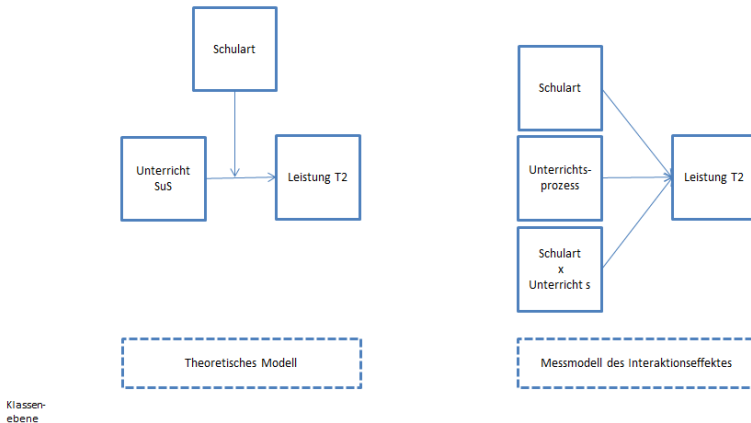


Abbildung 9. Vereinfachtes Modell eines Interaktionseffekts von Schulart und Unterrichtsprozessmerkmal auf Klassenebene.

Das linke Modell veranschaulicht die theoretische Annahme, das rechte Modell spezifiziert das analysierte Messmodell.

Angelehnt an Urban und Mayerl (2018).

Analysiert wird das theoretische Modell in Mplus 7 (Muthén und Muthén 1998–2013) als Regressionsanalyse mit der Schulart und dem Unterrichtsprozessmerkmal als Prädiktor. Zusätzlich werden die Interaktionsterme der Schulart mit dem Unterrichtsprozessmerkmal aufgenommen. Für die dummykodierte Schulartvariable mit dem Gymnasium als Referenzkategorie ergeben sich drei Interaktionsterme¹⁸¹⁹:

$$Y = \beta_0 + \beta_{UP}X + \beta_{RS}M_{RS} + \beta_{IGS}M_{IGS} + \beta_{MBG}M_{MBG} + \beta_{UPX} * \beta_{RS}M_{RS} + \beta_{UPX} * \beta_{IGS}M_{IGS} + \beta_{UPX} * \beta_{MBG}M_{MBG} + e_1 \quad (17)$$

Die relevante Information aus Gleichung (17) ist, dass insgesamt vier Haupteffekte vorliegen und drei Interaktionseffekte. X bezeichnet ein Unterrichtsprozessmerkmal mit dem zugehörigen Regressionsgewicht β_{UP} . M ist die Schulartvariable, die durch das Subskript jeweils eine der Schularten repräsentiert (RS=Realschule, IGS=integrierte Gesamtschule; MBG=Schule mit mehreren

18 Im Gegensatz zum vereinfachten Modell in Abbildung 9 werden hier alle Effekte der Schulart spezifiziert.

19 Die Darstellung der Moderationsanalyse erfolgt in Anlehnung an Jose (2013).

Bildungsgängen). Die zugehörigen Regressionsgewichte sind mit dem gleichen Subskript gekennzeichnet. Finden sich signifikante Effekte eines Interaktions-terms (Interaktionseffekte), so ist anzunehmen, dass für die jeweilige Schulart der Zusammenhang zwischen dem Unterrichtsprozessmerkmal und der Leistung (Y) zum zweiten Messzeitpunkt signifikant vom Zusammenhang im Gymnasium abweicht.

Bislang wurde das interessierende Moderations- bzw. Interaktionsmodell auf Unterrichtsebene dargestellt. Dieses Modell wird nun um die Individualebene erweitert, da nicht nur die Leistung zum zweiten Erhebungszeitpunkt analysiert werden soll, sondern die Leistungsentwicklung über zwei Messzeitpunkte. Hierfür wird auf Individualebene die Leistung zum ersten Erhebungszeitpunkt (T1) kontrolliert, d.h., es wird auf Individualebene eine Regression der Leistung zu T2 auf die Leistung zu T1 berechnet (siehe Abbildung 10). Das Mehrebenenmodell erlaubt es schließlich, das Intercept der Regression als „random effect“ zu spezifizieren, das somit zwischen den Schulklassen variiert (Hox 2013). In Abbildung 10 repräsentiert der ausgefüllte Kreis auf Klassenebene am Ende des Pfeiles von Leistung T1 zu Leistung T2 dieses „random intercept“. Das Intercept ist auf Klassenebene die abhängige Variable (Leistung T2 Intercept; in Anlehnung an Muthén und Muthén 2017). Somit bilden die bereinigten Klassenmittelwerte der Leistung zum zweiten Messzeitpunkt die abhängige Variable auf Klassenebene im Moderationsmodell. Auf Klassenebene wird, analog zu den Modellen in Fragestellung 2 und Fragestellung 3 für Antwortendenzen (Akquieszenz = ARS; Midpoint Responding = MPR) kontrolliert.

Abschließend sei noch kurz auf die Modellierung von Leistungsveränderung in den Analysen zu Fragestellung 4 eingegangen, da die Messung von Veränderung ein methodisch komplexes und kontrovers diskutiertes Thema darstellt (vgl. Campbell et al. 1999). Die beiden prominentesten Verfahren zur Analyse von Veränderung stellen der kovarianzanalytische Ansatz (wie in Abbildung 10 dargestellt) und die Analyse von Differenzwerten dar. In der methodischen Literatur besteht jedoch kein Konsens darüber, welches dieser beiden Verfahren zu bevorzugen ist (siehe bspw. Becker et al. 2006; Miller und Chapman 2001). In dieser Arbeit wird der kovarianzanalytische Ansatz verfolgt.

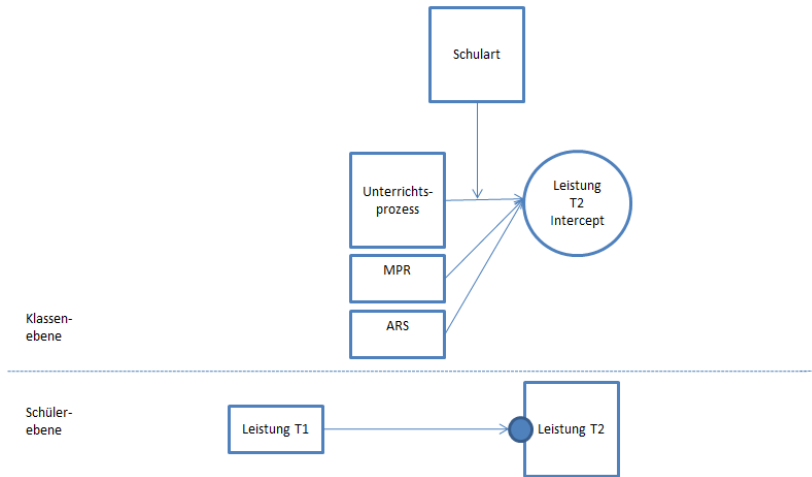


Abbildung 10. Vollständiges Mehrebenen-Moderationsmodell mit Kontrolle der Leistung auf Individualebene und Kontrolle von Antworttendenzen.
 Abkürzungen: (ARS=Akquieszenz, MPR=Midpoint Responding) auf Klassen-ebene.

7 Ergebnisse

In diesem Teil der Arbeit wird die Frage, welche Auswirkungen die Schulstruktur der Sekundarstufe I auf die Unterrichtsprozesse hat, empirisch für den Mathematikunterricht von Lernenden in der 9. bzw. 10. Klasse bearbeitet. Theoretisch wurden bereits Unterschiede in den Unterrichtsprozessen zwischen den Schularten herausgearbeitet (siehe Kapitel 2) und auch Vermittlungsmechanismen über die Lehrenden, die Schülergruppe und die Unterrichtsinhalte hergeleitet (Kapitel 3). Der Forschungsüberblick zu bisherigen Befunden aus Schulartvergleichen legte bereits Unterschiede in einzelnen Merkmalen zwischen den Schularten offen. Dennoch wurden Forschungslücken identifiziert (Abschnitt 3.5), die hier vertiefend bearbeitet werden.

Zunächst ist die Befundlage bisheriger Schulartvergleiche aus methodischer Sicht unsicher, da wichtige Vorannahmen bezüglich der Vergleichbarkeit der Daten nicht geprüft wurden. Abschnitt 7.1 widmet sich daher der Frage, inwieweit latente Konstrukte, welche Unterrichtsprozesse und damit zusammenhängende Merkmale beschreiben, invariant sind zwischen den Schularten und inwieweit die Daten aus Fragebogenstudien durch Antworttendenzen verzerrt sind. Hieraus leiten sich zum einen Schlussfolgerungen für die weiteren Analysen dieser Arbeit ab. Zum anderen werden die Ergebnisse zur Vergleichbarkeit der Daten auch hinsichtlich ihrer Bedeutung für die Interpretation von Befunden aus anderen Vergleichsstudien diskutiert.

Im nächsten Analyseteil (Abschnitt 7.2) wird der Kritik Rechnung getragen, dass umfassende Beschreibungen der Schularten bislang selten sind, da meist nur einzelne Merkmale vergleichend zwischen den Schularten betrachtet werden und häufig nur Aussagen über die Unterschiede zwischen Gymnasien und Hauptschulen oder zwischen Gymnasien und den anderen Schularten getroffen werden. Zudem wurden bislang keine Signifikanztests bzw. Effektstärkemaße für die Mittelwertdifferenzen berichtet, weshalb evtl. Unterschiede zwischen den Schularten überinterpretiert wurden. Diese Arbeit beschreibt daher die Schularten umfassend anhand der Schülergruppe, der Lehrenden, der Unterrichtsinhalte und der Unterrichtsprozesse. Die Merkmale wurden vorab auf ihre Vergleichbarkeit geprüft und entsprechende Folgerungen für die Beschreibung der Schularten getroffen. Zusätzlich wird die Varianz der Merkmale zwischen und innerhalb der Schularten berücksichtigt. Somit sind zuverlässige Beschreibungen der Schularten und ihre Kontrastierung möglich.

Abschnitt 7.3 geht schließlich auf die Frage ein, inwieweit die Unterschiede in den Unterrichtsprozessen über die schulartspezifischen Ausgangsbedingungen in der Zusammensetzung der Schülergruppe, den Überzeugungen der Lehrenden und den Unterrichtsinhalten vermittelt werden. Hierfür wurden für jene Merkmale aus den deskriptiven Analysen Mediationsmodelle berechnet, bei denen sich die Schulartunterschiede als signifikant erwiesen haben. Abschnitt 7.4 geht abschließend auf die Frage ein, inwieweit sich nicht nur die Unterrichtsprozesse, sondern auch deren Zusammenhänge mit der Leistungsentwicklung der Lernenden zwischen den Schularten unterscheiden.

Jeder Ergebnisteil endet mit einer zusammenfassenden Interpretation der Befunde und daraus resultierenden Schlussfolgerungen. Diese werden im Diskussionsteil tiefergehend in einem größeren Zusammenhang betrachtet.

7.1 Vergleichbarkeit von Daten zum Lehren und Lernen aus Fragebogenstudien

Inwieweit die in dieser Sekundäranalyse einbezogenen Skalen vergleichbar sind zwischen den Schularten, soll in diesem Abschnitt empirisch angegangen werden. Die inhaltliche Diskussion, ob und wie Vergleiche zwischen Gruppen, Ländern und Kulturen angestellt werden sollten, wurde in Kapitel 4 dargestellt. Es wurde gefolgert, Unterrichtsprozesse zwischen Schularten zu vergleichen und damit die Reflexion über Bildung und Bildungskonzepte zu stärken und Erkenntnisse über den Unterricht zu erweitern. Im vorliegenden Abschnitt soll die Frage nach der Vergleichbarkeit nun aus methodischer Sicht gestellt werden. Denn bei Gruppenvergleichen besteht die Gefahr, dass Unterschiede durch ein unpassendes Messinstrument erst erzeugt werden (Allemann-Ghionda 2004). Diese Tatsachen erschweren aus methodischer Sicht die vergleichende Betrachtung der Schularten, da sich Unterschiede zwischen den Schularten neben tatsächlichen Unterschieden als Artefakte erweisen könnten, welche durch die Erhebung selbst hervorgerufen werden (van de Vijver 1998).

Diese Arbeit prüft daher, ob die eingesetzten Messinstrumente die Konstrukte äquivalent und somit vergleichbar zwischen den Schularten erheben. Ein weiteres Anliegen ist es, zusätzlich zu den empirischen Analysen auch die Ursachen für mangelnde Vergleichbarkeit zu ergründen. Deshalb werden für Konstrukte, welche mangelnde Äquivalenz aufweisen, die Items inhaltlich analysiert. Zu-

dem wird geprüft, inwieweit die Daten durch Antworttendenzen verzerrt sind, und Schlussfolgerungen für die weiteren Analysen getroffen. Bei den Analysen wird aus methodischen Gründen unterschiedlich vorgegangen, je nachdem, ob Lehrende oder Lernende die Informationsquelle sind.

7.1.1 Äquivalenz der Messinstrumente

Konstrukt-, Methoden- und Itembias können den Vergleich von Daten bzw. die Äquivalenz der eingesetzten Messinstrumente beeinträchtigen. Je nachdem, welche Stufe der Invarianz vorliegt, können Daten nur innerhalb der Schularten analysiert werden oder es sind Vergleiche von Korrelationen und Mittelwerten zwischen den Schularten zulässig (siehe im Methodenteil Abschnitt 6.4). Ein gängiges Verfahren, um die Äquivalenz von Fragebogenskalen zwischen verschiedenen Gruppen zu prüfen, ist die mehrgruppenkonfirmatorische Faktorenanalyse (MGCFA). Diese bietet gut interpretierbare Statistiken, um die Messinvarianz zu beurteilen, und bietet sich an, um die Skalen der Lernenden zu testen. Für die Daten der Lehrenden ist hingegen die Fallzahl zu gering, um belastbare Ergebnisse aus MGCFA zu erhalten. Die Invarianz wird für die Skalen der Lehrenden daher mittels Faktorvergleichen aus Hauptkomponentenanalysen getestet.

7.1.1.1 Unterricht aus Sicht der Lernenden und Einstellungen der Lernenden

Die Invarianz der Messinstrumente wird in drei aufeinander aufbauenden Stufen getestet: konfigurale, metrische und skalare Invarianz. Tabelle 6 gibt die Fitindizes der einzelnen Modelle wieder (CFI, RMSEA, SRMR, BIC), auf deren Basis die Invarianz beurteilt werden kann. Die konfigurale Invarianz wird direkt über die absolute Höhe dieser Indizes beurteilt. Bei der metrischen Invarianz sowie der skalaren Invarianz werden hingegen die Unterschiede in der Modellpassung zwischen Modellen mit mehr bzw. weniger Restriktionen betrachtet. Das Modell mit mehr Restriktionen (höherer Stufe der Invarianz) darf nur unwesentlich schlechter sein als das Modell mit weniger Restriktionen (niedrigere Stufe der Invarianz).

In Tabelle 6 bedeutet ein negativer Wert eine schlechtere Passung des Modelles mit mehr Restriktionen bzw. höherer Invarianzstufe. Graue Zahlen indizieren,

dass der maximal akzeptierte Wert für konfigurale Invarianz bzw. der maximale Veränderungswert für metrische und skalare Invarianz überschritten ist und die entsprechende Invarianzstufe nicht erreicht wird. Die Indizes sind unterschiedlich sensibel gegenüber Verletzungen der Modellannahmen oder der Bestrafung von mehr frei geschätzten Parametern. Deshalb wird die Invarianz anhand mehrerer Indizes beurteilt. Die letzten drei Spalten fassen die Schlussfolgerungen für die drei Stufen der Invarianz zusammen. Liegt keine konfigurale Invarianz vor, können die Daten nur innerhalb der Schularten analysiert, aber keine Vergleiche zwischen Schularten vorgenommen werden. Nur wenn die Konstrukte metrisch invariant sind, dürfen korrelative Zusammenhänge zwischen den Schularten verglichen werden. Mittelwertvergleiche setzen zusätzlich skalare Invarianz voraus.

Aus Tabelle 6 lässt sich entnehmen, dass für 9 der 13 Konstrukte konfigurale Invarianz angenommen werden kann (✓). Für diese Konstrukte laden die Variablen in allen Gruppen auf dem jeweils theoretisch angenommenen Faktor, d.h., die verwendeten Items messen die Konstrukte vergleichbar zwischen den Gruppen. Für die vier Konstrukte, welche nicht konfigural invariant sind (x), wurden weitere Analysen durchgeführt, um mögliche Ursachen mangelnder Vergleichbarkeit herauszuarbeiten. Konfirmatorische Faktorenanalysen getrennt für die Gruppen indizieren, dass für die Konstrukte „kognitive Aktivierung“ sowie „Arbeitsmoral“ auch innerhalb der Gruppen keine ausreichende Modellpassung vorliegt. Hier scheint ein grundlegendes Problem des Messinstrumentes vorzuliegen. Vermutlich sind die Items nicht geeignet, um die intendierten Konstrukte zu messen. In der internationalen PISA-2012-Studie wurde die geringe Passung für Deutschland vermutlich nicht entdeckt. Zumindest im Technical Report (OECD 2014) werden lediglich die Reliabilitäten auf Länderebene – gemessen über Cronbachs α – angegeben, welche für Deutschland im akzeptablen Bereich liegen. Die Konstrukte „Einstellung zur Schule“ und „Einstellung gegenüber dem Lernen“ weisen innerhalb der Gruppen eine ausreichende Modellpassung auf. Ein Vergleich der Gruppenergebnisse zeigt allerdings Unterschiede in den korrelativen Zusammenhängen der Items zwischen den Schularten auf. Besonders die Interkorrelationen in der Gruppe der integrierten Gesamtschulen weichen stark von denen der anderen Schularten ab. Versuchsweise wurde die konfigurale Invarianz für die Konstrukte „Einstellung gegenüber dem Unterricht“ und „Einstellung zur Schule“ unter Ausschluss der integrierten Gesamtschule erneut getestet. Zwar erhöht sich so die Modellpassung deutlich, aber auch für die verbleibenden Gruppen ist die Modellpassung nicht ausreichend, um konfigurale Invarianz anzunehmen. Hieraus ist zu

Tabelle 6. Ergebnisse der Invarianzanalyse.

	Konfigural			Metrisch			Skalar			Konfig. invariant	Metrisch invariant	Skalar invariant
	CFI	RMSEA	SRMR	Δ CFI	Δ RMSEA	SRMR	Δ BIC	Δ CFI	Δ RMSEA	SRMR	Δ BIC	
Einstellung Schule	0,85	0,18	0,05	0,04	0,07	-0,01	42,02	-0,03	0,01	-0,01	35,38	x
Einstellung Lernen	0,85	0,17	0,04	0,05	0,08	-0,02	32,44	-0,29	-0,06	-0,05	-282,26	x
Instrumentelle Motivation	1,00	0,05	0,01	-0,01	-0,01	-0,04	31,55	-0,02	-0,02	-0,02	-21,14	x
Interesse	1,00	0,05	0,01	0,00	0,01	-0,03	54,08	0,00	0,01	0,00	61,41	✓
Arbeitsmoral	0,75	0,16	0,10	-0,02	0,01	0,01	9,36	-0,02	0,01	0,00	6,80	x
Klassenführung	1,00	0,02	0,01	0,00	0,02	-0,01	62,79	-0,01	-0,03	-0,02	12,00	✓
Kognitive Aktivierung	0,87	0,08	0,05	-0,01	0,01	-0,01	122,12	-0,02	0,00	-0,01	0,89	x
Konstruktive Unterstützung	1,00	0,04	0,02	0,00	0,01	-0,01	82,05	-0,01	-0,01	-0,01	18,51	✓
Formatives Assessment	0,98	0,07	0,03	0,00	0,02	-0,01	59,97	-0,02	-0,01	-0,01	14,95	✓
Schülerorientierte Unterrichtsmethoden	0,99	0,04	0,02	0,01	0,03	-0,01	60,82	-0,01	-0,03	-0,02	-2,36	✓
Lehrerorientierter Unterricht	0,95	0,05	0,05	0,00	0,01	-0,03	77,03	-0,01	0,00	-0,03	42,05	✓
Disziplin	1,00	0,06	0,02	0,00	0,01	-0,02	89,88	0,00	0,00	-0,01	136,50	✓
S-L-Beziehung	0,98	0,07	0,02	0,00	0,01	-0,01	76,69	-0,01	0,00	0,00	54,58	✓

Graue Zahlen indizieren, dass der maximal akzeptierte Wert für konfigurale Invarianz bzw. der maximale Veränderungswert für metrische und skalare Invarianz überschritten wird. Die letzten drei Spalten fassen die Ergebnisse für jede Invarianzstufe zusammen. Ein Haken (✓) indiziert, dass die entsprechende Invarianzstufe erreicht wird. Ein Kreuz (x) indiziert, dass die entsprechende Invarianzstufe nicht erreicht wird.

folgern, dass die integrierten Gesamtschulen bei diesen Konstrukten besonders stark von den anderen Schularten abweichen. Die Messung der Konstrukte scheint sich zwischen den anderen Schularten zwar ähnlicher zu sein, aber nicht ausreichend, um vergleichbare Konstruktrepräsentationen annehmen zu können.

Inhaltlich lässt sich die mangelnde Invarianz über die verwendeten Items begründen, welche die Konstrukte messen. Das Konstrukt „Einstellung gegenüber dem Lernen“ wird über vier Items erhoben. Zwei davon beziehen sich direkt auf den Beruf bzw. das Studium: *„Mich in der Schule anzustrengen wird mir den Zugang zum Studium ermöglichen“* bzw. *„Mich in der Schule anzustrengen wird mir helfen, eine interessante Arbeit zu finden.“* Aufgrund des selektiven Bildungssystems ist es auch für strebsame Lernende in bestimmten Bildungsgängen oder Schularten weniger aussichtsreich, ein Studium aufnehmen zu können. Wenn beispielsweise Lernende an Realschulen nicht zustimmen, dass das Lernen den Zugang zum Studium ermöglicht, muss dies nicht bedeuten, dass eine negative Einstellung gegenüber dem Lernen vorliegt. Eventuell drückt sich darin das Wissen aus, dass die Mittlere Reife keinen direkten Zugang zum Studium ermöglicht. Somit misst das Item *„Mich in der Schule anzustrengen wird mir den Zugang zum Studium ermöglichen“* für diese Lernende nicht die Einstellung gegenüber der Schule. Auf der anderen Seite mag es für Lernende, welche das Abitur anstreben, hilfreich sein, sich in der Schule anzustrengen, um ein Studium aufzunehmen. Über den Erfolg in den Berufseinstieg entscheiden aber häufig die Noten im Studium. Daher ist das Item *„Mich in der Schule anzustrengen wird mir helfen, eine interessante Arbeit zu finden.“* für diese Lernenden weniger relevant, um die Einstellung zum Lernen zu messen.

Die Items, welche die „Einstellung zur Schule“ messen sollen, sind bewertend hinsichtlich des Nutzens der Schule als Vorbereitung auf den Beruf, die Zukunft als Erwachsener und das erworbene Selbstvertrauen, bspw. *„Die Schule hat mich Dinge gelehrt, die für meinen zukünftigen Beruf nützlich sein könnten“*. Die Schularten haben jedoch verschiedene Zielsetzungen in der Bildung und der Erziehung, beispielsweise hat das Gymnasium den Fokus nicht auf der Berufsvorbereitung. Vielmehr haben die Unterrichtsinhalte einen viel stärkeren wissenschaftspropädeutischen Charakter als die anderer Schularten der Sekundarstufe. Zudem nimmt die kulturelle Bildung im Gymnasium einen wesentlichen Teil ein (siehe Abschnitt 2.2). Lernende an Gymnasien können das Beispielitem ablehnen und trotzdem eine positive Einstellung zur Schule haben.

Andererseits sind affektive Merkmale wie Selbstwert und Selbstwirksamkeit im Mittel an Gymnasien am höchsten ausgeprägt, an Hauptschulen am geringsten. Die anderen Schularten liegen meist dazwischen. Die Angabe nach dem erworbenen Selbstvertrauen erfasst somit nicht unbedingt die Einstellung gegenüber der Schule, sondern erfasst vielmehr die Ausprägung des Selbstvertrauens, welches sich nachweislich im Mittel zwischen den Schularten unterscheidet.

Zusammenfassend können die Operationalisierungen der Konstrukte „Einstellung zur Schule“ und „Einstellung gegenüber dem Lernen“ nicht als vergleichbar angesehen werden. Durch die mangelnde Invarianz und die inhaltlichen Schwierigkeiten bei den verwendeten Items können daher nur neun Konstrukte als konfigural invariant angenommen werden. Für diese Konstrukte wird im nächsten Schritt die metrische Invarianz geprüft. Die Ergebnisse zeigen, dass für alle neun Konstrukte gleiche Ladungen angenommen werden können und somit metrische Invarianz erreicht wird. Der Umgang mit der Antwortskala bezogen auf die Abstände der Antwortoptionen ist somit vermutlich vergleichbar zwischen den Lernenden der verschiedenen Schularten.

Die höchste Stufe der hier getesteten Invarianz, die skalare Invarianz, liegt für acht der untersuchten Konstrukte vor. Für das Konstrukt „instrumentelle Motivation“ ist der Skalenursprung hingegen nicht vergleichbar zwischen den Gruppen. Bei einer Häufigkeitsskala wäre beispielsweise der Ursprung identisch, wenn alle Befragten die Antwortoption „nie“ mit der Auftretenswahrscheinlichkeit von Null interpretieren. Für das Konstrukt „instrumentelle Motivation“ könnten Referenzgruppeneffekte ausschlaggebend sein. Die instrumentelle Motivation in Mathematik wird über eine vierstufige Antwortskala von „stimme völlig zu“ bis „stimme überhaupt nicht zu“ erhoben. Da hier objektive Kriterien fehlen, wird mit großer Wahrscheinlichkeit die eigene Einstellung im Vergleich zu den Klassenkameraden beurteilt (Clausen 2002). Allerdings bleibt unklar, warum der Referenzgruppeneffekt bei diesem Konstrukt stärker sein soll als bei anderen Konstrukten mit gleicher Antwortskala, bspw. bei dem Konstrukt „Interesse in Mathematik“.

Auch wenn die Gründe für fehlende Invarianz nicht gänzlich hergeleitet werden können, werden aus den empirischen Ergebnissen für die weiteren Analysen entsprechende Konsequenzen gezogen. Vier Konstrukte werden für die weiteren Analysen ausgeschlossen (Einstellung gegenüber der Schule, Einstellung gegenüber dem Lernen, Arbeitsmoral sowie kognitive Aktivierung). Alle ande-

ren Konstrukte eignen sich für korrelative Untersuchungen. Mittelwertvergleiche führen jedoch nur bei skalar invarianten Konstrukten zu verlässlichen Aussagen. Deshalb basieren die deskriptiven Beschreibungen der Schularten aus Sicht der Lernenden nur auf den acht Konstrukten, die skalare Invarianz aufzeigen.

7.1.1.2 Ziele im Mathematikunterricht und Potential zur kognitiven Aktivierung aus Sicht der Lehrenden

Die Invarianz der Messinstrumente wird über den Vergleich der Faktorstruktur aus Hauptkomponentenanalysen (PCA) und dem Kontingenzkoeffizient nach Tucker (1951) beurteilt. In einem ersten Schritt wird geprüft, ob die Items in allen Vergleichsgruppen auf den gleichen Faktoren laden. Ist dies gegeben, kann in Anlehnung an mehrgruppenkonfirmatorische Faktorenanalysen (MGCFA) Invarianz angenommen werden. Der Kontingenzkoeffizient gibt im zweiten Schritt an, ob das Ladungsmuster vergleichbar ist. Dieser Schritt prüft somit in Anlehnung an MGCFA auf metrische Invarianz. Die skalare Invarianz kann mit diesem Verfahren nicht getestet werden. Damit weicht die Invarianzanalyse der Lehrerskalen methodisch vom Vorgehen bei den Schülerskalen ab. Grund dafür ist die geringe Fallzahl bei den Lehrerdaten (siehe Abschnitt 6.2.2). Die Ergebnisse des Faktorstrukturvergleichs werden im Folgenden ausführlich und getrennt für die Ziele im Mathematikunterricht und das Potential zur kognitiven Aktivierung dargelegt.

Ziele im Mathematikunterricht

Die Ziele der Lehrenden im Mathematikunterricht wurden mit einer Skala erhoben, die bereits in der COACTIV-Studie entwickelt wurde. Dort wurden theoretisch und empirisch vier Zielbereiche bzw. Dimensionen unterschieden.

- Beherrschen von Routinen
- Modellierungsfähigkeit
- Anwendung im Alltag
- Interesse

Auch mit den PISA-Daten können diese vier Dimensionen unter Einbezug der Faktoreigenwerte (zwischen 4,55 und 1,17) und des Scree Plots in der PCA repliziert werden. Die Faktorstruktur weicht jedoch in einem Item ab (siehe

Tabelle 7). Item 10 („Interesse für die Anwendung der Mathematik“) lädt auf dem Faktor „Interesse“, statt auf dem Faktor „Anwendung im Alltag“. Inhaltlich ist dies gut nachvollziehbar. Die exakt gleiche Formulierung („Interesse für die Anwendung der Mathematik“) findet sich in Item 14. Dieses lädt übereinstimmend mit den Ergebnissen der COACTIV-Studie auf dem Faktor „Interesse“. Da es sich hier um Sekundäranalysen handelt, kann die Entscheidung ein Item doppelt zu erheben und verschiedenen Dimensionen zuzuordnen nicht geklärt werden. Für die weiteren Analysen werden hier beide Items der Dimension „Interesse“ zugeordnet.

Die relativ hohen Ladungen (zwischen 0,67 und 0,90) zeigen, dass die Variablen in hohem Ausmaß durch die jeweilige Zieldimension erklärt werden. Lediglich Item 5 („Schnelles Rechnen im Kopf bei Standardanwendungen“) lädt nur schwach auf der Zieldimension „Beherrschen von Routinen“ (0,41). Dies liegt vermutlich daran, dass die Lehrenden die Relevanz des Kopfrechnens sehr unterschiedlich beurteilen. Das Item weist mit einer Standardabweichung von $s=0,72$ die höchste Varianz in den Antworten auf. Gleichzeitig sind die Zusammenhänge mit den anderen Items vergleichsweise gering ($0,095 < r_{yx} < 0,385$). Inwieweit hier systematische Unterschiede zwischen den Schularten vorliegen, werden die folgenden Analysen getrennt nach Schulart zeigen.

Die Analysen innerhalb der Schularten zeigen, dass die Dimensionen „Anwendung im Alltag“ und „Modellierungsfähigkeit“ in allen Schularten mit den gleichen Items gemessen werden können (die zugehörigen Ergebnisse pro Schulart finden sich im Online-Anhang 12.4). Alle Items laden relativ hoch auf der Zieldimension. Leichte Abweichungen zeigen sich zwischen den Schularten bei der Dimension „Beherrschen von Routinen“. Das zugehörige Item 5 („Schnelles Rechnen im Kopf“), welches bereits in der Analyse über alle Schularten hinweg mit der großen Varianz aufgefallen ist, scheint zwischen den Schularten eine unterschiedliche Bedeutung zu haben. Bei Lehrenden an Schulen mit mehreren Bildungsgängen ist dieses Items stärker mit der Vermittlung von Interesse verbunden statt mit Routinen. Lehrende an Realschulen und integrierten Gesamtschulen verbinden schnelles Kopfrechnen eher mit Modellierungsfähigkeit. Für Lehrende an Gymnasien bildet dieses Item zusammen

Tabelle 7. Faktorstruktur der Ziele im Mathematikunterricht für die Gesamtstichprobe.

Item	Beherrschen von Routinen	Modellierungsfähigkeit	Anwendung im Alltag	Interesse
1 Sicherheit bei der Bearbeitung von Standardaufgaben	0,83	-0,01	0,00	0,10
2 Vermittlung von mathematischem Grundwissen	0,70	-0,08	-0,19	0,15
3 Sichere Beherrschung von mathematischen Routineoperationen	0,84	0,00	-0,17	0,15
4 Automatisierte Beherrschung der wichtigsten Algorithmen der Mittelstufe	0,70	0,29	-0,05	0,19
5 Schnelles Rechnen im Kopf bei Standardanwendungen	0,41	0,29	-0,31	0,29
6 Mathematisierungsfähigkeit entwickeln	0,04	0,76	-0,09	0,37
7 Alltagssituationen in mathematische Modelle zu übersetzen	0,12	0,69	-0,54	0,20
8 Mathematisch argumentieren können	0,08	0,80	-0,18	0,31
9 Die Angemessenheit eines mathematischen Modells erkennen können	0,08	0,78	-0,26	0,32
10 Interesse für die Anwendungen der Mathematik	0,18	0,22	-0,32	0,77
11 Mathematik im Alltag verstehen	0,17	0,22	-0,91	0,42
12 Mathematisches Wissen im Alltag anwenden können	0,14	0,27	-0,90	0,27
13 Freude am mathematischen Denken entwickeln	0,17	0,24	-0,25	0,78
14 Interesse für die Anwendungen der Mathematik	0,07	0,31	-0,21	0,70
15 Mathematik zu etwas persönlich Wichtigem machen	0,12	0,38	-0,23	0,68
16 Erreichen, dass die Schülerinnen und Schüler Spaß an der Mathematik haben	0,20	0,21	-0,14	0,70

mit Item 4 („Automatisierte Beherrschung der wichtigsten Algorithmen der Mittelstufe“) eine eigenständige Dimension. Für keine der Schularten ist Item 5 damit eindeutig der theoretischen Zieldimension zugeordnet.

Noch weniger eindeutig ist das Ladungsmuster der Dimension „Interesse“, welches innerhalb der Schularten nicht eindeutig replizierbar ist. In Realschulen und integrierten Gesamtschulen laden die Items entsprechend der theoretischen Zuordnung in COACTIV auf der Zieldimension, die Ladungen sind jedoch teils gering und weniger eindeutig. In Gymnasien ist das Ziel, Interesse zu vermitteln, stark mit der Anwendung im Alltag verbunden. In Schulen mit mehreren Bildungsgängen sind die Items, welche die Interessensvermittlung messen sollen, teilweise mit dem Ziel, Modellierungsfähigkeit zu vermitteln, verbunden. Inhaltlich ist dies nicht zu erklären. Aus methodischer Sicht hat eventuell die identische Formulierung von Item 10 und 14 zu Problemen geführt. Die Dimension „Interesse“ ist somit weder für die Gesamtgruppe, noch innerhalb der Schularten eindeutig besetzt. Konfigurale Invarianz ist hier nicht anzunehmen.

Die Dimension „Beherrschen von Routinen“ ist mit Ausnahme von Item 5 innerhalb der Schularten replizierbar. Item 5 („Schnelles Rechnen im Kopf bei Standardanwendungen“) scheint in Anbetracht der Zuordnung zu verschiedenen Zieldimensionen zwischen den Schularten unterschiedlich besetzt zu sein. Allerdings sind auch innerhalb der Schularten die Ladungen für dieses Item nicht immer eindeutig. Dieses Item wird damit wohl generell zwischen den Lehrenden nicht einheitlich interpretiert. Die Dimension „Interesse“ wird damit für die weiteren Analysen ausgeschlossen. Die Dimension „Beherrschen von Routinen“ wird unter Ausschluss von Item 5 als konfigural invariant betrachtet. Die finale PCA für die Dimensionen „Beherrschen von Routinen“, „Modellierungsfähigkeit“ und „Anwendung im Alltag“ zeigt für alle Zieldimensionen für das Modell über die Schularten hinweg (siehe Tabelle 8) und innerhalb der Schularten (siehe Online-Anhang 12.4.1) theoriekonforme Ladungen. Diese drei Dimensionen sind daher als konfigural invariant zu beurteilen.

Tabelle 8. Finale Dimensionen der Ziele im Mathematikunterricht für die Gesamtstichprobe.

Item	Beherrschen von Routinen	Modellierungsfähigkeit	Anwendung im Alltag
1 Sicherheit bei der Bearbeitung von Standardaufgaben	0,86	-0,02	0,10
2 Vermittlung von mathematischem Grundwissen	0,72	-0,14	-0,16
3 Sichere Beherrschung von mathematischen Routineoperationen	0,84	-0,07	-0,08
4 Automatisierte Beherrschung der wichtigsten Algorithmen der Mittelstufe	0,68	0,27	0,11
6 Mathematisierungsfähigkeit entwickeln	-0,02	0,79	0,08
7 Alltagssituationen in mathematische Modelle zu übersetzen	0,01	0,60	-0,38
8 Mathematisch argumentieren können	0,01	0,80	0,00
9 Die Angemessenheit eines mathematischen Modells erkennen können	0,00	0,78	-0,06
11 Mathematik im Alltag verstehen	0,02	0,01	-0,91
12 Mathematisches Wissen im Alltag anwenden können	0,00	0,05	-0,89

In einem zweiten Schritt wird zu der Zuordnung der Items zu den einzelnen Dimensionen auch die Höhe der Faktorladungen zwischen den Schularten verglichen. Dieser Schritt approximiert die metrische Invarianz, wie sie in mehrgruppenkonfirmatorischen Faktorenanalyse (MGCFA) getestet wird. Hinweise auf vergleichbare Ladungen liefert der Kontingenzkoeffizient nach Tucker (1951); ausführlicher beschrieben im Methodenteil in Abschnitt 6.4.2. Die Werte des Koeffizienten liegen mit $0,92 < C < 0,99$ deutlich über dem Grenzwert (siehe Tabelle 9). Somit wird für die Dimensionen „Beherrschen von Routinen“, „Modellierungsfähigkeit“ und „Anwendung im Alltag“ auch metrische Invarianz angenommen, weshalb korrelative Zusammenhänge zwischen den Schularten vergleichbar sind. Auf skalare Invarianz kann in der PCA nicht

getestet werden. Die Erfahrungen aus den MGCFA der Schülerskalen zeigen aber, dass konfigurale Invarianz, also die vergleichbare Erhebung der Konstrukte über die verwendeten Items, die größte Hürde bei der Vergleichbarkeit war. Unterschiede im Umgang mit der Antwortskala waren kaum gegeben. Für die weiteren Analysen werden die drei Dimensionen für die Ziele des Mathematikunterrichts daher auch für Mittelwertvergleiche herangezogen.

Tabelle 9. Kontingenzkoeffizienten für die Dimensionen der Ziele im Mathematikunterricht.

	C _{Routinen}	C _{Modellieren}	C _{Alltag}
MBG	0,98	0,98	0,94
RS	0,99	0,98	0,96
IGS	0,97	0,98	0,92
GY	0,98	0,95	0,94

Abkürzungen: MBG=Schulen mit mehreren Bildungsgängen, RS=Realschulen, IGS=integrierte Gesamtschulen, GY=Gymnasien.

Kognitiv herausfordernder Unterricht

Das Potential eines kognitiv herausfordernden Unterrichts wurde ebenfalls mit einer Skala erhoben, die bereits in der COACTIV-Studie²⁰ entwickelt wurde. Dort wurden theoretisch und empirisch drei Dimensionen zur kognitiven Aktivierung unterschieden:

- Kognitiv aktivierende Aufgaben,
- Insistieren auf Erklären und Begründen,
- Verständnisfördernde Variation von Aufgabenstellungen.

Auch mit den PISA-Daten²¹ können diese drei Dimensionen unter Einbezug der Faktoreigenwerte (zwischen 4,70 und 1,13) und des Scree Plots in der PCA repliziert werden. Wobei die Ladungen teilweise gering oder nicht eindeutig sind (siehe Tabelle 10). Besonders Item 3 (*„Ich lasse unterschiedliche Lösungswege von Aufgaben vergleichen“*) und Item 8 (*„Ich stelle auch Aufgaben,*

20 Vgl. <https://daqs.fachportal-paedagogik.de/search/show/survey/113>.

21 Vgl. <http://zib.education/de/forschung/abgeschlossene-projekte/sekundaeranalyse-der-pisa-plus-daten.html>.

bei denen man mehrere Lösungswege zeigen muss“) laden auf zwei Dimensionen relativ ähnlich (Item 3: $\lambda_1=0,41$ $\lambda_2=0,37$; Item 8: $\lambda_1=0,60$, $\lambda_2=0,51$). Beide Items betonen den Einsatz von Aufgaben und gleichzeitig kann mit der Aufgabe auch eine Begründung verbunden sein. Inwieweit hier systematische Unterschiede im Umgang mit diesen Items zwischen den Schularten vorliegen, werden die folgenden Analysen getrennt nach Schulart zeigen.

Innerhalb der Schularten findet sich die intendierte Faktorstruktur nicht eindeutig wieder. Besonders bei der Dimension „Kognitiv aktivierende Aufgaben“ finden sich Kreuzladungen. Die Items für diese Dimension wurden daher in einer separaten PCA tiefergehend analysiert. In allen Schularten ergab sich für die Dimension „Kognitiv aktivierende Aufgaben“ eine 2-Faktorenlösung. Allerdings verteilten sich die Items zwischen den Schularten völlig unterschiedlich auf die zwei Faktoren, weshalb nicht davon auszugehen ist, dass es sich hierbei um ein vergleichbares Konstrukt handelt. Die Dimension „Kognitiv aktivierende Aufgaben“ wird daher aus den weiteren Analysen ausgeschlossen und eine weitere PCA für die verbleibenden Dimensionen „Insistieren auf Erklären und Begründen“ und „Verständnisfördernde Variation von Aufgabenstellungen“ berechnet. Für die verbleibenden Dimensionen ist die Faktorstruktur über die Schularten hinweg und innerhalb der Schularten eindeutig und entsprechend den theoretischen Annahmen aus der COACTIV-Studie. Diese beiden Dimensionen werden daher in Anlehnung an die mehrgruppenkonfirmatorischen Faktorenanalyse als konfigural invariant beurteilt.

Tabelle 10. Faktorstruktur des Potentials zur kognitiven Aktivierung für die Gesamtstichprobe.

	Kognitiv aktivierende Aufgaben	Insistieren auf Erklären und Begründen	Verständnis- fördernde Variation von Aufgaben- inhalten
1 Ich lasse Aufgaben bearbeiten, für die es keinen sofort erkennbaren Lösungsweg gibt.	0,78	-0,14	-0,11
2 Ich stelle Fragen, die man nicht spontan beantworten kann, sondern die zum Nachdenken zwingen.	0,63	0,16	0,01
3 Ich lasse unterschiedliche Lösungswege von Aufgaben vergleichen und bewerten.	0,41	0,37	-0,01
4 Ich stelle Aufgaben, bei denen es nicht allein auf das Rechnen, sondern vor allem auf den richtigen Ansatz ankommt.	0,66	-0,05	0,19
5 Ich stelle Aufgaben, bei denen es auf das Verständnis des mathematischen Gedankengangs ankommt.	0,70	-0,10	0,26
6 Ich stelle Aufgaben, die keine eindeutige Lösung haben, und lasse dies erklären.	0,76	0,06	-0,06
7 Ich stelle Aufgaben, für deren Lösung man Zeit zum Nachdenken braucht.	0,72	0,03	0,06
8 Ich stelle auch Aufgaben, bei denen man mehrere Lösungswege zeigen muss.	0,48	0,47	-0,13
9 Ich halte Schüler/-innen dazu an, ihre Gedankengänge genau zu erklären.	0,11	0,68	0,08
10 Ich verlange, dass Schüler/-innen ihre Arbeitsschritte ausführlich begründen.	-0,18	0,85	0,04
11 Bei mir wissen die Schüler/-innen, dass sie ihre Aussage immer begründen müssen, auch wenn ich es nicht extra dazusage.	0,00	0,70	0,16

12	In jeder Klassenarbeit bringe ich mindestens eine Aufgabe, in der eine Begründung gefordert ist.	0,10	0,03	0,59
13	Ich wechsle die Einkleidung von Aufgaben so, dass man erkennt, ob die mathematische Idee wirklich verstanden wurde.	0,05	0,13	0,81
14	Ich variiere Aufgaben so, dass man sieht, was die Schüler/-innen verstanden haben.	-0,08	0,00	0,85

Tabelle 11. Finale Dimensionen des Potentials zur kognitiven Aktivierung für die Gesamtstichprobe.

		Kognitiv aktivierende Aufgaben	Insistieren auf Erklären und Begründen
9	Ich halte Schüler/-innen dazu an, ihre Gedankengänge genau zu erklären.	0,68	0,14
10	Ich verlange, dass Schüler/-innen ihre Arbeitsschritte ausführlich begründen.	0,89	-0,11
11	Bei mir wissen die Schüler/-innen, dass sie ihre Aussage immer begründen müssen, auch wenn ich es nicht extra dazusage.	0,79	0,03
12	In jeder Klassenarbeit bringe ich mindestens eine Aufgabe, in der eine Begründung gefordert ist.	0,15	0,53
13	Ich wechsle die Einkleidung von Aufgaben so, dass man erkennt, ob die mathematische Idee wirklich verstanden wurde.	0,03	0,90
14	Ich variiere Aufgaben so, dass man sieht, was die Schüler/-innen verstanden haben.	-0,12	0,91

In einem zweiten Schritt wird neben der Zuordnung der Items zu den einzelnen Dimensionen auch die Höhe der Faktorladungen zwischen den Schularten verglichen. Dieser Schritt approximiert die metrische Invarianz, wie sie in MGCFA getestet wird. Hinweise auf vergleichbare Ladungen liefert der Kontingenzkoeffizient nach Tucker (1951); ausführlicher beschrieben im Metho-

denteil in Abschnitt 6.4.2. Die Werte des Koeffizienten liegen mit $0,88 < C < 1,00$ deutlich über dem Grenzwert (siehe Tabelle 12). Somit wird für die Dimensionen „Insistieren auf Erklären und Begründen“ und „Verständnisfördernde Variation von Aufgabenstellungen“ metrische Invarianz angenommen, weshalb korrelative Zusammenhänge zwischen den Schularten vergleichbar sind. Auf skalare Invarianz kann in der PCA nicht getestet werden. Die Erfahrungen aus der mehrgruppenkonfirmatorischen Faktorenanalyse der Schülerskalen zeigen aber, dass konfigurale Varianz, also die vergleichbare Erhebung der Konstrukte über die verwendeten Items, die größte Hürde bei der Vergleichbarkeit ist. Unterschiede im Umgang mit der Antwortskala sind kaum gegeben. Für die weiteren Analysen werden die zwei Dimensionen kognitiv herausfordernden Unterrichts auch für Mittelwertvergleiche herangezogen.

Tabelle 12. Kontingenzkoeffizienten für die Dimensionen des Potentials zur kognitiven Aktivierung.

	C _{Begründen}	C _{Variation}
MBG	0,99	0,99
RS	0,99	1,00
IGS	0,98	0,99
GY	0,86	0,88

Abkürzungen: MBG=Schulen mit mehreren Bildungsgängen, RS=Realschulen, IGS=integrierte Gesamtschulen, GY=Gymnasien.

7.1.2 Antworttendenzen und deren Effekte auf Mittelwerte

In standardisierten Befragungen stellen Verzerrungen durch Antworttendenzen eine besondere Herausforderung dar, vor allem bei Items im Likert-Format (Bogner und Landrock 2015). Likert-Items werden meist eingesetzt, wenn Einstellungen oder Verhaltensweisen erfragt werden. In der Schul- und Unterrichtsforschung werden bspw. Einstellungen zur Schule und dem Lernen, Unterrichtsprozesse oder Klimavariablen über Fragebogenitem im Likert-Format erhoben. Antworttendenzen können als eine Form des Methodenbias die Invarianz dieser Messinstrumente gefährden (siehe bspw. Cheung und Rensvold 2000; Grimm und Church 1999). Nicht immer kann eine Invarianzanalyse jedoch die Verzerrung der Daten durch Antworttendenzen aufdecken. So zeigen Simulationsstudien, dass nur bei extremen Verzerrungen mangelnde Invarianz durch Antworttendenzen aufgedeckt werden kann (Liu et al. 2017). Es emp-

fehlt sich daher, einzelne Formen der Antworttendenzen zusätzlich zu prüfen. Daher werden die Likert-Skalen in dieser Arbeit zusätzlich auf Antworttendenzen geprüft. Ziel ist es, das Ausmaß der Antworttendenzen der Lernenden sowie der Lehrenden zu ergründen. Zudem interessiert, wie stark die Gruppenmittelwerte und die Gruppenvergleiche durch Antworttendenzen verzerrt sind.

In untenstehender Tabelle 13 sind die klassengemittelten Antworttendenzen der Lernenden sowie die Antworttendenzen der zugehörigen Lehrenden gemittelt pro Schulart dargestellt. Ein Akquieszenzwert (ARS) von 1 bedeutet eine völlige Zustimmung zu allen Likert-Items, ein Wert von -1 bedeutet die generelle Ablehnung der Aussagen über alle einbezogenen Items hinweg. Die Tendenz mittlere Antworten zu wählen (MPR), wird in einem Wertebereich zwischen 0 und 1 angegeben. Ein Wert von 0 bedeutet die ausschließliche Wahl der mittleren Antwortkategorien, ein Wert von 1 bedeutet hingegen die Wahl der Extremkategorien (siehe im Methodenteil Abschnitt 6.4.3).

Tabelle 13. Tendenz zur Zustimmung und zur Wahl von mittleren Kategorien pro Schulart und gesamt.

	Lernende		Lehrende	
	ARS	MPR	ARS	MPR
MBG	0,19	0,38	0,27	0,38
RS	0,17	0,41	0,24	0,39
IGS	0,20	0,42	0,21	0,36
GY	0,14	0,39	0,21	0,36
η^2 Schulart	0,07	0,05	0,02	0,01

Abkürzungen: MBG=Schulen mit mehreren Bildungsgängen, RS=Realschulen, IGS=integrierte Gesamtschulen, GY=Gymnasien, ARS=Akquieszenz, MPR=Mid-Point-Responding; $\eta^2 \geq 0,06$ ist fett gedruckt.

Aus der Tabelle ist ersichtlich, dass sich für alle Befragten Antworttendenzen zeigen. In allen Schularten haben Lernende wie Lehrende die Tendenz, den Aussagen zuzustimmen und die mittleren Antwortkategorien zu bevorzugen. Die Tendenz zur Mitte ist bei Lehrenden wie Lernenden ähnlich, akquieszentes Verhalten zeigen die Lehrenden stärker als die Lernenden. Wobei hier nicht fälschlicherweise gefolgert werden darf, dass Lehrende generell eine größere Tendenz zu akquieszentem Antwortverhalten haben als Lernende. Da die Fragebögen unterschiedliche Inhalte abfragen, kann dies auch auf situationsbezogene Faktoren zurückgeführt werden, welche die Auftretenswahrscheinlichkeit

von Antworttendenzen ebenfalls bedingen (ausführlicher im Theorieteil in Abschnitt 4.2). Es kann jedoch übereinstimmend mit bisherigen Befunden gefolgert werden, dass Daten, welche mittels Items im Likert-Format erhoben wurden, durch Antworttendenzen verzerrt sind.

Für den Vergleich zwischen den Schularten lässt sich aus den Ergebnissen folgern, dass die Ausprägungen der Antworttendenzen vor allem bei den Lernenden unterschiedlich ausfallen. Für die Akquieszenz zeigt sich ein mittlerer Effekt der Schulart ($\eta^2=0,07$). An integrierten Gesamtschulen und Schulen mit mehreren Bildungsgängen findet sich die größte Zustimmungstendenz. Die Zustimmungstendenz in Gymnasien und Realschulen ist vergleichsweise gering. Für die Tendenz, mittlere Antwortkategorien zu wählen, zeigt sich ein geringerer Effekt der Schulart ($\eta^2=0,05$). Die stärkste Tendenz, mittlere Antwortkategorien zu wählen, findet sich bei den Lernenden in Klassen an Schulen mit mehreren Bildungsgängen. Die geringste Tendenz findet sich bei Lernenden in Klassen an Realschulen und integrierten Gesamtschulen. Weniger theoriekonform ist das Ergebnis bei Lernenden an Gymnasium, welche eine ähnlich hohe Tendenz zur Mitte zeigen wie Lernende an Schulen mit mehreren Bildungsgängen. Bei den Lehrenden zeigen sich für beide Antworttendenzen nur marginale Effekte der Schulart ($\eta^2=0,02$ bzw. $\eta^2=0,01$). Lehrende an Schulen mit mehreren Bildungsgängen tendieren geringfügig stärker zu akquieszentem Antwortverhalten als Lehrende der anderen Schularten. Tendenzen, mittlere Antwortkategorien zu wählen, treten geringfügig stärker bei Lehrenden an Realschulen auf im Vergleich zu Lehrenden an integrierten Gesamtschulen und Gymnasien.

Die Ergebnisse stimmen mit der Hypothese überein, dass die Tendenz, Antworten zuzustimmen, als Strategie bei kognitiver Überforderung auftritt. Lernende an Schulen mit mehreren Bildungsgängen und integrierten Gesamtschulen, welche die höchste Zustimmungstendenz zeigen, weisen im Durchschnitt die geringsten kognitiven Fähigkeiten auf, verglichen mit den Lernenden an Realschulen und Gymnasien. Auch für die Tendenz, mittlere Antwortkategorien zu wählen, scheint es einen Zusammenhang mit den kognitiven Fähigkeiten zu geben. Lediglich die vergleichsweise hohe Tendenz bei Lernenden an Gymnasien, widerspricht dem Zusammenhang von kognitiven Fähigkeiten und der Tendenz, mittlere Antwortkategorien zu wählen. Lehrende hingegen, welche sich hinsichtlich der kognitiven Fähigkeiten wahrscheinlich nicht systematisch zwischen den Schularten unterscheiden, zeigen auch über die Schularten hin-

weg hypothesenkonform ähnliche Tendenzen zu akquieszentem Verhalten und die Tendenz zur Mitte.

Um das Ausmaß der Verzerrung durch Antworttendenzen einschätzen zu können, interessiert im nächsten Schritt, inwieweit sich die Interpretation der Daten unterscheidet, wenn Antworttendenzen bei der Datenanalyse von Mittelwertunterschieden berücksichtigt werden. Nachfolgende Tabellen stellen die Ergebnisse von Varianzanalysen mit und ohne Kontrolle von Antworttendenzen gegenüber (siehe im Methodenteil Abschnitt 6.7). Die jeweils erste Spalte für ein Merkmal gibt den Mittelwert pro Schulart ohne Kontrolle von Antworttendenzen an und den zugehörigen Effekt der Schulart (η^2). In der zweiten Spalte ist die Abweichung in den Ausprägungen angegeben, wenn Antworttendenzen kontrolliert werden. Ein positiver Wert gibt an, dass sich die Ausprägung unter Kontrolle der Antworttendenzen erhöht, umgekehrtes gilt bei einem negativen Wert. Die Ausprägung für den Effekt der Schulart (η^2) ist **fett** markiert, wenn sich die Einschätzung der Schulartunterschiede durch die Berücksichtigung der Antworttendenzen verändert, bspw. wenn ein ehemals geringer Effekt der Schulart ($\eta^2 < 0,6$) unter Kontrolle der Antworttendenzen als mittlerer Effekt der Schulart einzuschätzen ist. Mittelwerte sind **fett** markiert, wenn sich unter Kontrolle der Antworttendenzen das Differenzmuster zwischen den Schularten bedeutend ändert. Deutlicher werden Unterschiede in den Differenzmustern in der graphischen Gegenüberstellung. Hierfür sei auf den Online-Anhang 12.5 verweisen. In untenstehenden Tabellen sind auch jene Skalen aufgeführt, für welche keine skalare Invarianz vorliegt und somit Mittelwertvergleiche nicht zulässig sind (siehe Abschnitt 7.1.1). Um mangelnde Invarianz zu verdeutlichen sind die Werte in **grau** dargestellt. Die Unterschiede der Skalen, welche nicht skalar invariant sind werden hier nicht tiefergehend interpretiert, da Mittelwertvergleiche nicht zulässig sind. Diese und ähnliche Merkmale wurden und werden jedoch in anderen Analysen interpretiert, weshalb hier die Unterschiede, welche sich mit und ohne Kontrolle von Antworttendenzen ergeben, zumindest dokumentiert werden. Offensichtlich ist, dass sich für diese Skalen die Differenzmuster ändern, wenn Antworttendenzen kontrolliert werden.

Bei den invarianten Skalen, für welche Mittelwertvergleiche zulässig sind, ändern sich über die Merkmale hinweg die z-standardisierten Mittelwerte zwischen $0 \leq \Delta \leq 0,87$ Einheiten, in positiver wie auch in negativer Richtung. Der Effekt der Schulart (η^2) weist Differenzen zwischen $0 \leq \Delta \leq 0,05$ Einheiten auf, ebenfalls in positiver wie negativer Richtung. Entscheidender als die absoluten

Veränderungen ist jedoch, ob sich inhaltlich Unterschiede in der Interpretation ergeben. Eine Änderung in η^2 kann entscheidend sein, ob der Effekt der Schulart und somit Schulartunterschiede als gering eingestuft werden, oder aber mittlere Effekte vorliegen. Für das Antwortmuster ist bedeutend, ob sich die Abstände zwischen den Mittelwertunterschieden ändern, oder sich sogar die Reihenfolge der Schularten in den Ausprägungen ändert.

Für die dreizehn Merkmale, welche Mittelwertvergleiche erlauben, ändert sich die Interpretation des Schularteffekts (η^2) und somit der Bedeutung der Mittelwertunterschiede in sechs Fällen. Für das Ziel, Mathematisierungsfähigkeiten zu vermitteln, sowie für die kognitive Aktivierung durch Variation der Aufgaben ist der Effekt der Schulart unter Kontrolle von Antworttendenzen als mittelgroß einzustufen. Ohne Kontrolle der Antworttendenzen läge nur ein geringer Effekt vor. Für die anderen vier Merkmale – Anwendung im Alltag, Schüler-Lehrer-Beziehung, Lehrerorientierter Unterricht und konstruktive Unterstützung – werden ohne Kontrolle von Antworttendenzen die Schulartunterschiede in den aufgeführten Merkmalen hingegen überschätzt. Tatsächlich liegt nur ein geringer Effekt der Schulart vor statt eines mittleren Effekts.

Veränderung in Mittelwerten und Effektstärken (η^2) unter Kontrolle von Antworttendenzen.

Abkürzungen: Ohne AT: Ausprägung ohne Kontrolle von Antworttendenzen, Δ mit AT=Veränderung unter Kontrolle von Antworttendenzen; Bedeutende Veränderungen sind fett gedruckt. Invariante Skalen sind grau markiert.

Tabelle 14. Veränderung der Werte bei Überzeugungsmerkmalen der Lehrenden unter Kontrolle von Antworttendenzen.

	Anwendung im Alltag		Mathe-matisieren		Basiswissen		Interesse	
	Ohne AT	Δ mit AT	Ohne AT	Δ mit AT	Ohne AT	Δ mit AT	Ohne AT	Δ mit AT
MBG	0,43	-0,07	-0,43	-0,10	0,32	-0,06	0,02	-0,12
RS	0,10	0,01	-0,14	0,00	0,13	0,03	0,01	0,00
IGS	-0,21	0,02	-0,04	0,04	0,18	0,02	-0,15	0,04
GY	-0,16	0,02	0,16	0,03	-0,06	0,00	-0,07	0,04
η^2 Schulart	0,06	-0,01	0,05	0,03	0,02	0,00	0,00	0,00

Tabelle 15. Veränderung der Werte bei affektiven Schülermerkmalen unter Kontrolle von Antworttendenzen.

	Einstellung Schule		Einstellung Lernen		Arbeitsmoral		Motivation		Interesse		S-L-Beziehung	
	Ohne AT	Δ mit AT	Ohne AT	Δ mit AT	Ohne AT	Δ mit AT	Ohne AT	Δ mit AT	Ohne AT	Δ mit AT	Ohne AT	Δ mit AT
MBG	0,24	-0,01	-0,10	0,01	-0,04	-0,05	0,08	-0,05	0,05	-0,06	-0,02	0,00
RS	0,10	-0,01	-0,03	0,00	0,00	-0,02	0,01	-0,01	-0,03	-0,01	0,16	-0,02
IGS	0,16	-0,09	0,08	-0,07	0,11	-0,14	0,15	-0,10	-0,03	-0,11	0,07	-0,09
GY	-0,12	0,04	0,04	0,03	-0,02	0,07	-0,05	0,05	-0,02	0,06	-0,06	0,04
n ² Schulart	0,14	-0,03	0,03	0,01	0,01	0,01	0,04	-0,04	0,01	0,03	0,06	-0,01

Tabelle 16. Veränderung der Werte bei Unterrichtsmerkmalen unter Kontrolle von Antworttendenzen.

	Formatives Assessment		Lehrer-orientierter Unterricht		Schüler-orientierter Unterricht		konstruktive Unterstützung		kognitive Aktivierung		kognitive Aktivierung durch Begründungen		kognitive Aktivierung durch Variation		kognitive Aktivierung durch Aufgaben		Klassenführung		Disziplin	
	Ohne AT	Δ mit AT	Ohne AT	Δ mit AT	Ohne AT	Δ mit AT	Ohne AT	Δ mit AT	Ohne AT	Δ mit AT	Ohne AT	Δ mit AT	Ohne AT	Δ mit AT	Ohne AT	Δ mit AT	Ohne AT	Δ mit AT	Ohne AT	Δ mit AT
MBG	0,35	-0,06	0,19	-0,06	0,36	-0,07	0,30	-0,09	0,00	-0,07	-0,34	-0,04	0,26	0,00	-0,28	-0,03	0,09	-0,08	-0,01	-0,87
RS	0,11	-0,01	-0,04	-0,01	0,03	0,00	0,10	-0,01	-0,02	-0,01	-0,33	0,01	-0,32	0,01	-0,03	-0,01	0,00	-0,01	-0,04	0,00
IGS	0,24	-0,11	0,18	-0,12	0,31	-0,09	0,06	-0,15	0,20	-0,13	-0,29	0,02	-0,31	0,00	0,11	0,01	0,01	-0,13	-0,04	-0,07
GY	-0,17	0,06	-0,05	0,06	-0,13	-0,62	-0,14	0,08	0,00	0,07	0,08	-0,02	0,15	-0,01	0,07	0,01	0,03	0,07	-0,07	0,18
n ² Schulart	0,17	-0,05	0,06	-0,04	0,10	-0,04	0,07	-0,03	0,02	0,00	0,04	0,02	0,05	0,01	0,02	0,01	0,00	0,01	0,01	0,02

Differenzmuster ändern sich bei sechs der dreizehn Merkmale, wobei dies ausschließlich Merkmale betrifft, welche durch die Lernenden eingeschätzt wurden. Die *Schüler-Lehrer-Beziehung* wird mit Kontrolle von Antworttendenzen homogener eingeschätzt. Nur die Klassenmittelwerte an Realschulen weichen in positiver Richtung von den anderen Schularten ab. Ohne Kontrolle der Antworttendenzen zeigen sich für die Schüler-Lehrer-Beziehung erkennbare Unterschiede, wie sie auch in anderen Analysen gefunden werden. An Gymnasien wird die Schüler-Lehrer-Beziehung ohne Kontrolle von Antworttendenzen (fälschlicherweise) am geringsten eingeschätzt, an Realschulen und integrierten Gesamtschulen am höchsten.

Auch für das Merkmale *Interesse am Mathematikunterricht* ändert sich das Differenzmuster und zugleich die Rangfolge der Schularten. Ohne Kontrolle von Antworttendenzen sind die mittleren Ausprägungen in den Schularten nahezu gleich, gering weichen Schulen mit mehreren Bildungsgängen ab, deren mittlere Ausprägung in diesem Merkmal geringfügig höher ist. Unter Kontrolle von Antworttendenzen werden die Abstände zwischen den mittleren Ausprägungen der Schularten deutlicher. Klassen an Gymnasien weisen dann das höchste mittlere Interesse auf, Schulen mit mehreren Bildungsgängen leicht niedrigere Werte als Gymnasien. Die stärksten Änderungen zwischen den Ergebnissen mit und ohne Kontrolle von Antworttendenzen ergeben sich für die integrierten Gesamtschulen. Werden Antworttendenzen berücksichtigt, haben Klassen an integrierten Gesamtschulen mit deutlichem Abstand das geringste Interesse für Mathematik.

Ebenso für die Unterrichtsprozessmerkmale *Klassenführung* und *Disziplin* ändert sich das Differenzmuster. Unter Kontrolle von Antworttendenzen sind die Differenzen zwischen den Schularten deutlich größer. Bedeutender ist jedoch die geänderte Rangfolge der Schularten. Beim Merkmal Disziplin ist diese sogar spiegelbildlich und besonders für Klassen an Schulen mit mehreren Bildungsgängen sind die mittleren Ausprägungen deutlich niedriger als ohne Kontrolle der Antworttendenzen.

In der Einschätzung des *Lehrerorientierten Unterrichts* bilden sich ohne Kontrolle von Antworttendenzen zwei gegensätzliche Gruppen. Klassen an Schulen mit mehreren Bildungsgängen und an integrierten Gesamtschulen haben im Mittel ähnlich hohe Werte. Die andere Gruppe formiert sich aus Klassen an Gymnasien und Realschulen, welche im Mittel negativere Werte bei diesem

Merkmal haben. Werden Antworttendenzen kontrolliert, ergibt sich ein differenzierteres Bild mit einer unterscheidbaren Rangfolge der Schularten. Die höchsten Werte haben im Mittel Klassen an Schulen mit mehreren Bildungsgängen, gefolgt von Klassen an integrierten Gesamtschulen. Die dritthöchsten Werte haben im Mittel Klassen an Gymnasien. Klassen an Realschulen haben im Mittel die niedrigsten Ausprägungen bei diesem Merkmal.

Für das Merkmal *konstruktive Unterstützung* finden sich die größten Veränderungen in den Mittelwerten der Schularten. Dennoch ändert sich das Differenzmuster für Schulen mit mehreren Bildungsgängen, Realschulen und Gymnasien nur unwesentlich. Klassen an integrierten Gesamtschulen weisen jedoch unter Kontrolle von Antworttendenzen im Mittel deutlich negativere Werte bei diesem Merkmal auf und haben somit unter Kontrolle von Antworttendenzen im Mittel die negativsten Ausprägungen. Ohne Kontrolle von Antworttendenzen liegen die mittleren Ausprägungen für die integrierte Gesamtschule im Mittelfeld.

7.1.3 Zusammenfassung der Ergebnisse zur Vergleichbarkeit und Schlussfolgerungen für die weiteren Analysen

Die Analysen in diesem Ergebnisteil gingen der Frage nach, inwieweit die in dieser Sekundäranalyse einbezogenen Skalen vergleichbar sind zwischen den Schularten. Mangelnde Vergleichbarkeit kann zum einen systembedingt sein, da sich aus der historischen Entwicklung heraus mit den spezifischen Zielen der Schularten Möglichkeiten und Grenzen für die inhaltliche Gestaltung des Unterrichts ergeben und sich auch die Bildungs- und Berufsperspektiven der Lernenden an verschiedenen Schularten unterscheiden. Zum anderen verteilen sich die Lehrenden und Lernenden nicht zufällig auf die Schularten, weshalb individuelle Merkmale mit den interessierenden Merkmalen der Schularten konfundiert sind.

Die Invarianzanalysen der Schüler- und Lehrerskalen haben gezeigt, dass die vergleichbare Erfassung von Konstrukten über Fragebogenitems durchaus eine Herausforderung darstellt. Bei den Lernenden waren 4 von 13 Skalen, bei den Lehrenden 2 von 7 Skalen nicht konfigural invariant. Somit kann für diese Skalen nicht davon ausgegangen werden, dass die Items das Konstrukt vergleichbar in den Gruppen erfassen. Vergleiche zwischen den Schularten sind

für diese Konstrukte nicht zulässig. Bis auf ein Konstrukt liegt für alle Konstrukte, welche konfigural invariant sind, auch metrische und skalare Invarianz vor. Der Umgang mit der Antwortskala scheint daher nahezu gleich zwischen den Gruppen.

Inhaltliche Analysen der Items liefern Hinweise auf die Ursache mangelnder konfiguraler Invarianz. Für die Skalen der Lernenden konnte herausgearbeitet werden, dass durch Unterschiede in den Zugangschancen, welche sich durch die verschiedenen Schularten ergeben, sowie unterschiedliche Bildungsziele der Schularten die inhaltliche Bedeutung einzelner Items zwischen den Schularten abweicht. Bei den Lehrerskalen hat sich gezeigt, dass die abgefragten Methoden und Ziele zwischen den Schularten teils verschieden interpretiert werden. Bei den teils mehrdeutig formulierten Items scheint sich auch die Schwerpunktlegung zwischen den Schularten zu unterscheiden. Dies könnte unter anderem auf Unterschiede in der Ausbildung der Lehrpersonen und Unterschieden in den Bildungszielen der Schularten zurückgeführt werden.

Über die Invarianzanalysen hinaus wurde zusätzlich untersucht, ob die Daten aus der Fragebogenerhebung systematisch durch Antworttendenzen verzerrt sind. Die Tendenzen, Fragen zuzustimmen sowie mittlere Antwortkategorien zu wählen, konnten für Lernende wie auch für Lehrende nachgewiesen werden. Systematische schulartspezifische Unterschiede im Ausmaß der Verzerrung liegen jedoch nur für die Lernenden vor. Diese stimmen mit der Theorie überein, dass die Wahrscheinlichkeit von Antworttendenzen steigt, je niedriger die kognitiven Fähigkeiten sind.

Für die Lehrenden hat sich gezeigt, dass die Antworttendenzen die Differenzmuster der Mittelwertunterschiede zwischen den Schularten erwartungsgemäß nicht wesentlich verändern. Allerdings hat sich bei zwei der drei Skalen zu den Zielen des Mathematikunterrichts die Interpretation des Schularteffektes unter Kontrolle von Antworttendenzen geändert. Somit ist davon auszugehen, dass Antworttendenzen die Varianz der gegebenen Antworten beeinflussen. Dies kann sich jedoch in zweierlei Weisen auswirken. So wurde unter Kontrolle von Antworttendenzen bei einer Skala der Effekt der Schulart stärker, bei der anderen Skala geringer. Bei den Lernenden verzerren Antworttendenzen sowohl die Differenzmuster als auch die Interpretation des Schularteffektes. Die Auswirkungen sind jedoch sehr unterschiedlich. Antworttendenzen können sowohl zu homogeneren als auch heterogeneren mittleren Einschätzungen zwischen den

Schularten führen. Nicht immer, aber häufig ändert sich die Rangfolge der Schularten. Auch der Effekt der Schulart kann, ähnlich wie bei den Lehrerskalen, stärker oder schwächer ausfallen, wenn Antworttendenzen berücksichtigt werden.

Welche Konsequenzen sind nun aus den Ergebnissen zur Vergleichbarkeit der Daten für die hier vorliegende Arbeit zu ziehen? Für die hier vorliegende Arbeit werden resultierend aus den Ergebnissen alle weiteren Analysen unter Kontrolle von Antworttendenzen durchgeführt. Dies betrifft Lehrer- wie auch Schülerskalen. Nur so können Mittelwerte, Effektstärken und Zusammenhänge zwischen den Merkmalen korrekt analysiert und interpretiert werden. Darüber hinaus werden die Schularten deskriptiv anhand von Mittelwerten nur mit skalar invarianten Konstrukten inhaltlich beschrieben. Konstrukte, welche eine geringere Invarianzstufe aufweisen, werden im deskriptiven Vergleich zwar aufgeführt, aber entsprechend als nicht vergleichbar markiert und folglich inhaltlich nicht interpretiert. Sie werden jedoch berücksichtigt, um an die Ergebnisse anderer Schulartvergleiche anzuschließen. Analog werden Zusammenhänge nur für ausgewählte Konstrukte analysiert, welche mindestens metrische Invarianz aufweisen.

7.2 Deskription der Schularten

Im theoretischen Teil dieser Arbeit wurden die Schularten vergleichend beschrieben und gegenübergestellt. In diesem Teil der Studie wird der Frage nach den Auswirkungen des gegliederten Bildungssystems der Sekundarstufe I auf das Lehren und Lernen in den Klassen deskriptiv nachgegangen und die Befunde mit den Erkenntnissen aus der bisherigen vergleichenden Lehr-Lern-Forschung (siehe den Forschungsüberblick in Abschnitt 3.5) verknüpft. Hierfür werden Merkmale der Schülergruppe, der Überzeugungen von Lehrenden, der Inhalte des Mathematikunterrichts sowie Unterrichtsprozessmerkmale im Fach Mathematik vergleichend zwischen den Schularten betrachtet.

Ziel dieser Arbeit ist es, die Prozesse im Unterricht zu beschreiben, deshalb interessieren hier nicht die Merkmale auf individueller Ebene, sondern die Klassenmittelwerte. Folglich werden in der Deskription die Klassenmittelwerte gemittelt pro Schulart berichtet (siehe Methodenteil Abschnitt 6.5). Auf Seiten der Lernenden werden die gemittelten Eigenschaften und Verhaltensweisen

daher als kognitives, sozioökonomisches und normatives Lernumfeld (im Theorieteil Abschnitt 3.4.3) bezeichnet. Weitere Merkmale der Beschreibung sind die Überzeugungen der Lehrenden zu den Zielen des Mathematikunterrichts (Abschnitt 7.2.5), die Unterrichtsinhalte (Abschnitt 7.2.6) sowie die Prozesse des Mathematikunterrichts (Abschnitt 7.2.7). Für alle Merkmale wird die Varianz in den Klassenmittelwerten innerhalb der Schularten berücksichtigt, indem das Effektstärkemaß η^2 bei der Interpretation der Ergebnisse berücksichtigt wird. Nur bei einem $\eta^2 \geq 0,06$ kann von einem „mittleren Effekt“ der Schulart und somit von bedeutenden Unterschieden zwischen den Schularten ausgegangen werden. In den vorangegangenen Analysen wurde gezeigt, dass die Daten, welche mittels Likert-Skalen erhoben wurden, durch Antworttendenzen verzerrt werden. Die hier beschriebenen Mittelwerte wurden daher unter Kontrolle von Antworttendenzen berechnet.

Ziel der deskriptiven Analyse ist es, Unterschiede im Lehren und Lernen zwischen den Schularten – Schulen mit mehreren Bildungsgängen (MBG), integrierte Gesamtschule (IGS), Realschule (RS) und Gymnasium (GY) – zu identifizieren. Zudem wird reflektiert, in welchen Merkmalen die Ergebnisse von anderen Forschungsbefunden abweichen und inwieweit dies durch Antworttendenzen (bei Likert-Skalen) und die Varianz der Klassen innerhalb der Schularten bedingt sein könnte. Für das kognitive und soziokulturelle Lernumfeld ist die Befundlage unstrittig, weshalb hier nur kurz die Ergebnisse in den bisherigen Forschungsstand eingeordnet werden. Für die anderen Merkmale wird im Anschluss an die Beschreibung der Ergebnisse dieser Arbeit ausführlich der Bezug zu den Ergebnissen anderer Forschungsarbeiten (wie in Abschnitt 3.5 zusammengefasst) hergestellt. Der abschließende Absatz 7.2.8 fasst die Ergebnisse der deskriptiven Befunde und deren Reflexion kurz zusammen. Hier wird auch ein Überblick der Merkmale gegeben, die bedeutende Schulartunterschiede beschreiben.

7.2.1 Zusammensetzung der Schülergruppe

Die Lernenden, welche sich innerhalb einer Klasse zusammenfinden, prägen mit ihren Eigenschaften, Fähigkeiten und Verhaltensweisen das Lernumfeld des Unterrichts. Im deutschen Bildungssystem zeigt sich, dass die auf Schülerleistungen basierende Allokation der Lernenden an verschiedene Schularten sowie soziale Selektionsprozesse zu einer schulartspezifischen Zusammensetzung der

Schülerschaft führen (Klinge 2016; Maaz 2017). Wie genau sich die Klassen in den Schularten bezogen auf die Leistung, den sozioökonomischen Hintergrund sowie bezüglich affektiver Merkmale im Mittel zwischen den Schularten unterscheiden, zeigen die folgenden Beschreibungen zum kognitiven, sozioökonomischen und normativen Lernumfeld.

7.2.2 Kognitives Lernumfeld

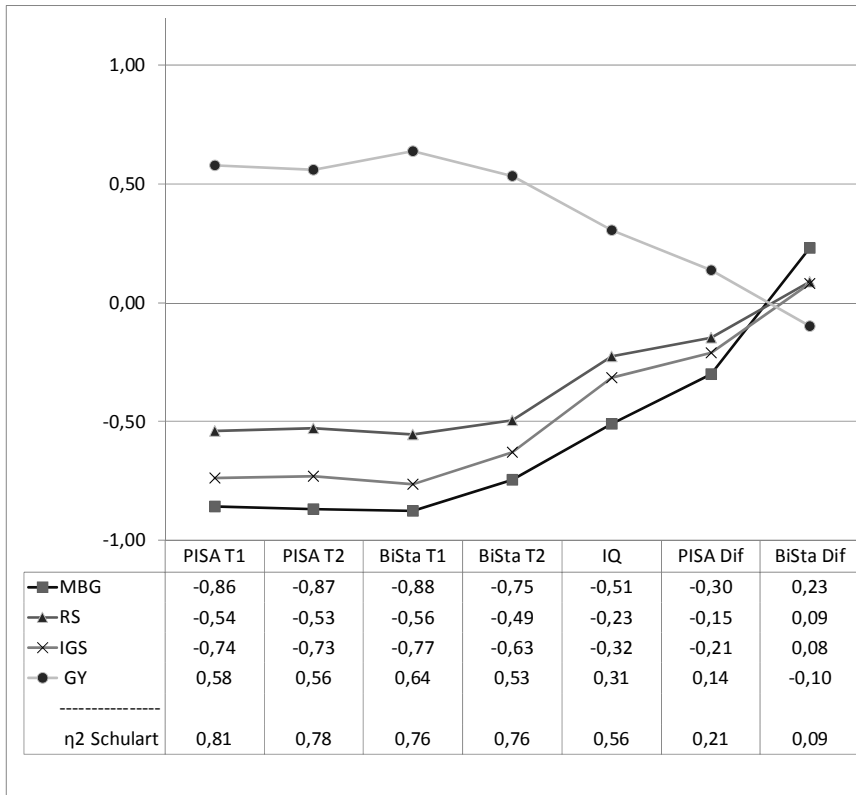


Abbildung 11. Mittlere Zusammensetzung der Klassen nach kognitiven Merkmalen in den Schularten.

Abkürzungen: MBG=Schulen mit mehreren Bildungsgängen, RS=Realschulen, IGS=integrierte Gesamtschulen, GY=Gymnasien.

Das kognitive Lernumfeld in den Schulklassen wird hier beschrieben über die mittlere Klassenleistung in Mathematik sowie den durchschnittlichen Intelligenzquotient (IQ) der Klassen. Die Kompetenzen in Mathematik werden über zwei unterschiedliche Tests erfasst: Die Mathematikaufgaben in PISA messen die mathematischen Fähigkeiten entsprechend des Literacy-Konzepts, d.h., die Fähigkeiten, „Mathematik in zahlreichen Kontexten anzuwenden, zu interpretieren und Formeln anzuwenden“ (Sälzer, Reiss et al. 2013, S. 51). Die Mathematikaufgaben des Bildungstrends erfassen hingegen curriculumsbezogene Fähigkeiten, welche in den Bildungsstandards (BiSta) formuliert sind (Ehmke et al. 2017). Für die Lernenden werden die Klassenleistungen der ersten Erhebung im Jahr 2012 (PISA T1 und BiSta T1) sowie der zweiten Erhebung im Jahr 2013 berichtet (PISA T2 und BiSta T2). Zudem wird die Leistungsentwicklung gemessen über die Differenz der Schülerleistungen zwischen den Messzeitpunkten (PISA Dif und BiSta Dif). Der IQ wurde mit dem BEFKI (siehe Methodenteil Abschnitt 6.2.1) nur in der ersten Erhebung im Jahr 2012 erfasst.

Bei den mathematischen Leistungen zeigen sich für beide Leistungstests zu beiden Zeitpunkten die erwarteten Mittelwertdifferenzen zwischen den Schularten, wie sie auch in anderen Studien gefunden werden (z. B. Baumert et al. 2004; Baumert, Lehmann, Lehrke, Schmitz, Clausen, Hosenfeld, Köller et al. 1997; Guill et al. 2010; Kunter et al. 2005; Schiepe-Tiska et al. 2013; Wendt et al. 2010). Klassen an Gymnasien verzeichnen im Mittel die höchsten Leistungswerte. Mit großem Abstand folgen die Realschulen und integrierten Gesamtschulen. Die geringsten Leistungswerte liegen im Mittel in allen Tests für die Klassen an Schulen mit mehreren Bildungsgängen vor. Der Effekt der Schulart ist mit $\eta^2 > 0,76$ sehr hoch. Die Differenzen zwischen den Schularten sind daher stark auf die Schulformzugehörigkeit zurückzuführen. Werden nicht die mathematischen Fähigkeiten sondern der IQ herangezogen, zeigt sich die gleiche Rangfolge. Jedoch sind die Unterschiede zwischen den Schularten deutlich geringer. Auch der Effekt der Schulart ist mit $\eta^2 = 0,56$ etwas niedriger als bei den Leistungstests, aber immer noch als sehr hoch zu beurteilen.

Die Leistungsentwicklung gemessen mit dem PISA-Test entspricht ebenfalls den Ergebnissen aus anderen Studien, die einen Schereneffekt in der Entwicklung zwischen Gymnasien und anderen Schularten identifizieren (bspw. Baumert et al. 2006b; Guill et al. 2010; Köller 1996; Scharenberg 2012). In Gymnasialklassen verzeichnen die Lernenden auch in der hier vorliegenden Studie

im Mittel die größten Leistungszugewinne. Mit größerem Abstand folgen die Realschulen und integrierten Gesamtschulen. Die geringste Leistungsentwicklung wird im Mittel für Klassen an Schulen mit mehreren Bildungsgängen gemessen. Der Effekt der Schulart ist mit $\eta^2=0,21$ als sehr hoch einzuschätzen. Die Schularten unterscheiden sich daher deutlich in der Leistungsentwicklung gemessen mit dem PISA-Test. Interessante Abweichungen ergeben sich, wenn die Leistungsentwicklung mit dem BiSta-Test erhoben wird. Für diese curriculumsnahen Tests ergeben sich spiegelbildliche Differenzen in der Leistungsentwicklung zwischen den Schularten. Lernende in Klassen an Schulen mit mehreren Bildungsgängen verzeichnen im Mittel die größte Leistungsentwicklung. Realschulen und integrierte Gesamtschulen liegen im Mittelfeld – mit sehr ähnlichen Schulartmittelwerten. Die geringsten Leistungsentwicklungen liegen für Lernende in Klassen an Gymnasien vor. Der Effekt der Schulart ist mit $\eta^2=0,09$ sichtlich geringer als bei der Leistungsentwicklung gemessen mit dem PISA-Test. Trotzdem können bei diesem mittleren Effekt bedeutende Unterschiede zwischen den Schularten angenommen werden.

Damit liegen zwei konträre Ergebnisse vor, deren genaueren Erklärung es bedarf. Die Leistungsentwicklungen gemessen mit dem PISA-Test stimmen mit Befunden anderer Studien überein, welche für die Entwicklung der Mathematikleistung in der Sekundarstufe einen Schereneffekt identifizieren können (Baumert et al. 2006b; bspw. Becker et al. 2006). Die Leistungsentwicklung gemessen mit dem BiSta-Test ist hingegen konform mit Erkenntnissen aus der KESS-Studie (siehe Abschnitt 3.5.1). Dort konnte zwar der Schereneffekt für die Mathematikleistung von der 4. zur 6. Klasse und von der 7. zur 8. Klasse gefunden werden, nicht jedoch von der 8. zur 10. Klasse, weshalb die These aufgestellt wurde, dass die Differenzen in der Leistungsentwicklung in der Sekundarstufe I nicht kontinuierlich über alle Jahre hinweg vorliegen. Die hier gefundenen Ergebnisse lassen jedoch vermuten, dass nicht nur die Klassenstufe sondern auch die Testinhalte relevant bei der Beschreibung von Leistungsentwicklungen sind. In dieser Arbeit ist es nun erstmals möglich, anhand einer Stichprobe die Leistungsentwicklungen im Fach Mathematik anhand verschiedener Mathematiktests mit unterschiedlichen Schwerpunkten gemeinsam zu betrachten.

Mögliche Erklärungen für die unterschiedlichen Ergebnisse zwischen den Leistungstests lassen sich auf methodische wie auch auf inhaltliche Aspekte zurückführen. Auf der methodischen Seite sind systematische Stichprobenausfälle zu

verzeichnen, welche mit der Testleistung zum ersten Erhebungszeitpunkt im Jahr 2012 sowie mit herkunftsspezifischen Merkmalen zusammenhängen (Heine et al. 2017). Da die Anzahl der Lernenden, welche den BiSta-Test bearbeitet haben, deutlich geringer ist als die Anzahl der Lernenden, welche den PISA-Test bearbeitet haben ($n_{\text{BiSta}}=1900$ $n_{\text{PISA}}=4600$), wirken sich die Stichprobenausfälle bei dem BiSta-Test vermutlich stärker aus als beim PISA-Test. Ob dies die Daten in solcher Größenordnung verzerrt, ist jedoch fraglich.

Plausibler ist die inhaltliche Erklärung für die Unterschiede in den Ergebnissen zwischen den Leistungstests. Die Bildungsstandards haben in den Ländern einen vergleichsweise engen Bezug zu den Lehrplänen und Curricula der Bundesländer (Ehmke et al. 2017) und somit sehr wahrscheinlich auch zu den realisierten Lerngelegenheiten im Unterricht. Somit ist der BiSta-Test sensibler als der PISA-Test für das schulisch erworbene Wissen (Naumann 2019). Unter dieser Annahme spiegelt sich in den Ergebnissen vermutlich ein verstärkter Wissenserwerb im Zuge der Vorbereitungen auf anstehende Abschlussprüfungen wider, wie die Mittlere Reife und den erweiterten Hauptschulabschluss. Durch den anstehenden Wechsel in die Berufswelt oder in die Sekundarstufe II sind Lernende an Realschulen, integrierten Gesamtschulen und Schulen mit mehreren Bildungsgängen in besonderem Maße angehalten, gute schulische Leistungen zu erbringen, da die Abschlussnoten den weiteren Bildungsweg maßgeblich beeinflussen. Ein vergleichbarer Anreiz ist an den Gymnasien in der zehnten Klasse nicht gegeben. Stattdessen scheinen die Lernenden in Klassen an Gymnasien durch den Fokus dieser Schulart auf reflektierendes und analytisches Denken besser auf die im PISA-Test erhobenen Inhalte vorbereitet zu werden. Ergebnisse aus dem Teilprojekt SteG-A der Studie zur Entwicklung von Ganztagschulen untermauern die Hypothese, dass Lernende am Übergang in die Berufswelt bzw. in die weiterführende Schule ihre Lernaktivitäten verstärkt auf die schulischen Fachinhalte fokussieren. Eine längsschnittliche Analyse der Teilnahme von Realschülern an außerschulischen Aktivitäten zeigt, dass in der neunten und zehnten Klasse die Aktivitäten in Ganztagsangeboten im Vergleich zu den Vorjahren deutlich geringer werden. Nur bei den fachbezogenen Angeboten ist ein deutlicher Anstieg in der Teilnahme zu verzeichnen (SteG-Konsortium 2016, S. 37). Die Erklärung über steigende Lernaktivitäten in der Abschlussphase in Realschulen, integrierten Gesamtschulen und Schulen mit mehreren Bildungsgängen ist auch plausibel für ähnliche Befunde bzgl. der Leistungsentwicklungen in Mathematik in der KESS-Studie. Dort wurden ebenfalls curriculumsbezogene Tests für die Leistungsmessung in Mathematik

herangezogen (Guill et al. 2010). Wie oben beschrieben, konnte zwar der bekannte Schereneffekt für die Mathematikleistung von der 4. zur 6. Klasse und von der 7. zur 8. Klasse gefunden werden, nicht jedoch von der 8. zur 10. Klasse.

Das Differenzmuster des kognitiven Lernumfeldes deckt sich somit mit Ergebnissen, wie sie auch in anderen Studien gefunden werden. Es wird deutlich, dass trotz der Leistungsstreuung innerhalb der Klassen und Schulen (siehe Baumert et al. 2006b) die Leistung und Leistungsentwicklung in erheblichem Maße auf die Schulart zurückzuführen ist. Für die Leistungsentwicklung zeigt sich zum einen der vielfach replizierte Schereneffekt bezogen auf die in PISA erhobene Fähigkeit, Mathematik in zahlreichen Kontexten anzuwenden und zu interpretieren sowie Formeln anzuwenden. Spiegelbildlich ist die Leistungsentwicklung hingegen bei den in den Bildungsstandards definierten und curriculumsnahen Mathematikfähigkeiten. Somit ist anzunehmen, dass die Differenzen in der Leistungsentwicklung nicht nur über die Klassenstufen hinweg variieren, sondern auch zwischen verschiedenen Bereichen der mathematischen Fähigkeiten.

7.2.3 Sozioökonomisches Lernumfeld

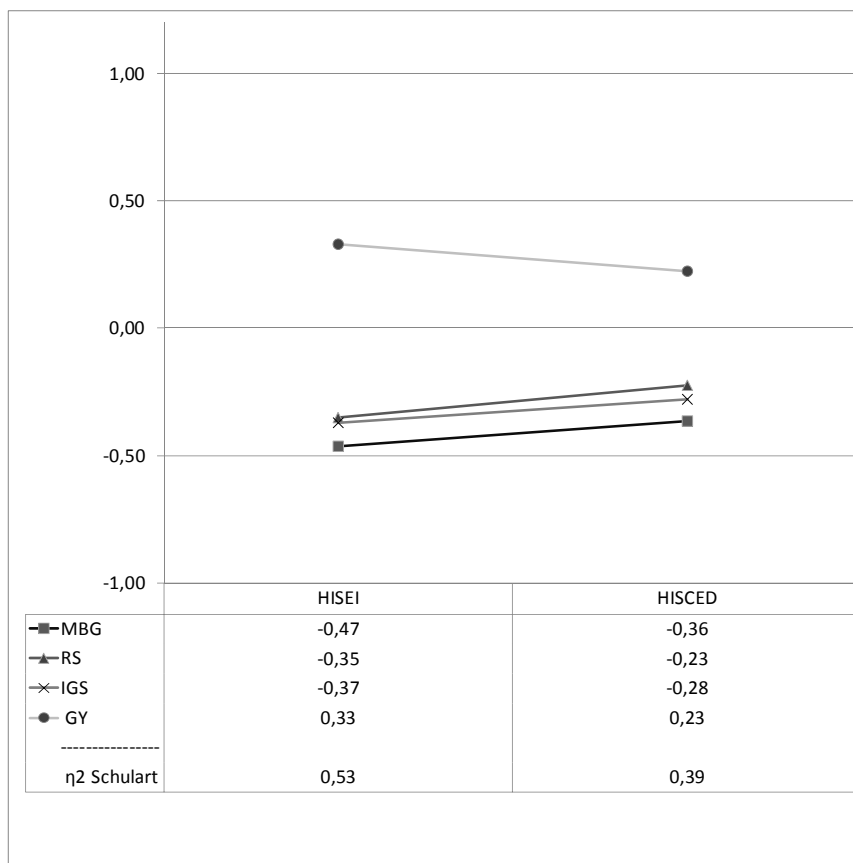


Abbildung 12. Mittlere Zusammensetzung der Klassen nach sozioökonomischen Merkmalen in den Schularten.

Abkürzungen: MBG=Schulen mit mehreren Bildungsgängen, RS=Realschulen, IGS=integrierte Gesamtschulen, GY=Gymnasien.

Das sozioökonomische Lernumfeld wird erhoben über den Berufsstatus der Eltern (HISEI) und den höchsten Abschluss der Eltern (HISCED). Der Berufsstatus approximiert vor allem die finanzielle Situation in den Familien. Dieser korreliert in Deutschland meist hoch mit dem Schulabschluss, welcher unter anderem das Unterstützungspotential und Bildungsaspirationen im Elternhaus approximieren soll. Der höchste Schulabschluss (HISCED) wird hier dichotom

gemessen, indem unterschieden wird, ob ein Elternteil mindestens Abitur hat oder nicht. Der HISEI ist hingegen ein kontinuierliches Maß (siehe im Methodenteil Abschnitt 6.2.1).

Für den HISEI und den HISCED zeigen sich erwartungsgemäß bedeutende Unterschiede in den Klassenmittelwerten besonders im Vergleich vom Gymnasium zu den anderen Schularten. In Klassen an Gymnasien haben Lernende erwartungsgemäß im Mittel Elternteile mit dem höchsten Berufsstatus und den größten Anteil an Eltern mit Abitur oder einem höheren Abschluss. Mit größerem Abstand folgen Realschulen und integrierte Gesamtschulen. Die Lernenden in Klassen an diesen Schularten sind sich in den mittleren Ausprägungen dieser Merkmale sehr ähnlich. Lernende in Klassen an mehreren Bildungsgängen haben im Mittel Elternteile mit dem geringsten Berufsstatus und weniger Elternteile mit Abitur oder einem höheren Abschluss als Lernende in Klassen der anderen Schularten. Der Effekt der Schulart ist bei dem Berufsstatus der Eltern sowie bei dem höchsten Abschluss mit $\eta^2=0,53$ und $\eta^2=0,39$ sehr hoch. Übereinstimmend mit zahlreichen Befunden zur Verteilung der Lernenden auf die verschiedenen Schularten (bspw. Baumert et al. 2006b; Prenzel et al. 2013) zeigt sich hier deutlich die Segregation im Verteilungsprozess auf die Schularten (vgl. Maaz et al. 2010).

7.2.4 Normatives Lernumfeld

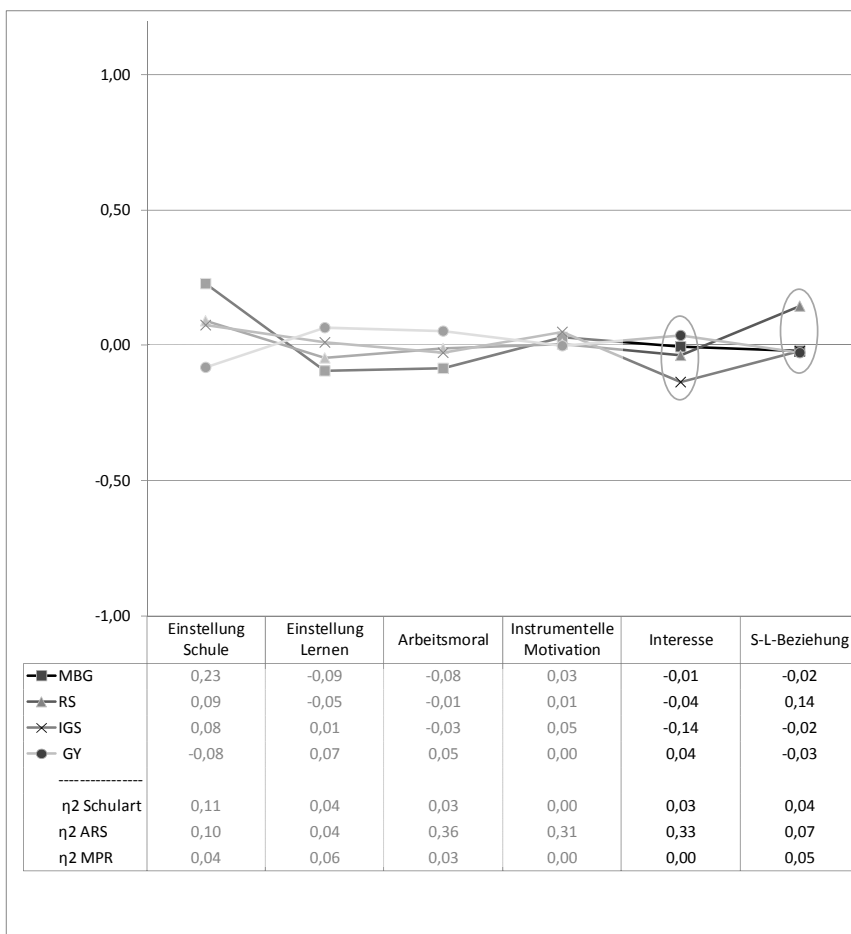


Abbildung 13. Mittlere Zusammensetzung der Klassen nach affektiven Merkmalen in den Schularten.

Konstrukte, welche nicht skalar invariant sind, sind grau unterlegt. Mittelwertvergleiche sind für diese Konstrukte nicht zulässig.

Abkürzungen: MBG=Schulen mit mehreren Bildungsgängen, RS=Realschulen, IGS=integrierte Gesamtschulen, GY=Gymnasien, ARS=Akquieszenz, MPR=Mid-Point-Responding.

Das normative Lernumfeld wird hier nur über zwei Skalen beschrieben: das Interesse für Mathematik sowie die Schüler-Lehrer-Beziehung. Diese Merkmale werden im Folgenden anhand der Differenzmuster und der Effektstärken interpretiert. Für die anderen vier Skalen kann nicht davon ausgegangen werden, dass die Konstrukte auf skalarem Niveau vergleichbar zwischen den Schularten gemessen werden. Somit sind Mittelwertvergleiche nicht zulässig (siehe Abschnitt 7.1.1). Diese vier Merkmale sind in Abbildung 13 aufgeführt und grau unterlegt. Sie werden im Folgenden jedoch nicht vergleichend zwischen den Schularten interpretiert.

Das Interesse der Lernenden für Mathematik ist im Mittel an den Klassen in Gymnasien am höchsten. Geringfügig niedriger ist das mittlere Interesse an Klassen in Realschulen und Schulen mit mehreren Bildungsgängen. Klassen an integrierten Gesamtschulen weichen stärker von den anderen Schularten ab. Hier ist das Interesse für Mathematik am geringsten. Die Differenzen in den Mittelwerten sind zwischen den Schularten jedoch nur marginal und die Unterschiede nicht als bedeutend zu bewerten, denn der Effekt der Schulart ist mit $\eta^2=0,03$ nur gering. Ähnlich verhält es sich beim Merkmal Schüler-Lehrer-Beziehung. Die Klassen in Schulen mit mehreren Bildungsgängen, integrierten Gesamtschulen und Gymnasien unterscheiden sich im Mittel nicht bzw. kaum. Lediglich die Klassen an Realschulen weisen im Mittel höhere Werte auf. Der Effekt der Schulart ist mit $\eta^2=0,04$ nur gering, weshalb keine bedeutenden Unterschiede zwischen den Schularten vorliegen.

Die hier beschriebenen Mittelwerte wurden unter Kontrolle von Antworttendenzen berechnet und die Interpretation bezieht die Varianz in den Merkmalen über das Effektstärkemaß η^2 ein. Für beide hier beschriebenen Merkmale hat sich in der Analyse zu den Antworttendenzen herausgestellt, dass sich unter Kontrolle der Antworttendenzen die Differenzmuster ändern (siehe Abschnitt 7.1.2). Es interessiert daher, inwieweit die Ergebnisse dieser Studie von den Ergebnissen in anderen Studien abweichen und ob dies auf die abweichende Berechnung und Interpretation zurückgeführt werden kann.

Für das Merkmal Interesse in Mathematik weichen die Ergebnisse in dieser Arbeit von Erkenntnissen aus anderen Studien ab (Baumert, Lehmann, Lehrke, Schmitz, Clausen, Hosenfeld, Köller 1997; vgl. Becker und Neumann 2016). Zwar ist das mittlere Interesse für Mathematik an Gymnasien übereinstimmend mit anderen Studien am höchsten. Allerdings sind die Unterschiede zu den

anderen Schularten gering. Lediglich Klassen an integrierten Gesamtschulen weisen im Mittel ein sichtlich niedrigeres Interesse für Mathematik auf. Solch große mittlere Abweichungen vom Gymnasium wurden bislang nur für Hauptschulen beschrieben, nicht für integrierte Gesamtschulen (siehe Abschnitt 3.5.1). Der stark negative Wert für die integrierten Gesamtschulen wird auch nur deutlich, wenn die Mittelwerte unter Kontrolle von Antworttendenzen berechnet werden (siehe Tabelle 15). Trotz der starken Differenzen von integrierten Gesamtschulen zu den anderen untersuchten Schularten sind die tatsächlichen Unterschiede als marginal einzustufen. Vermutlich findet sich hier deutlich mehr Varianz im untersuchten Merkmal auf Individual und Klassen-ebene als auf Ebene der Schulart. Somit ist unter Einbezug der Varianz und der Kontrolle von Antworttendenzen nicht davon auszugehen, dass sich die Klassen in ihrem Interesse für Mathematik bedeutend zwischen den Schularten unterscheiden. Im Gegensatz zur Interpretation in anderen Studien (Baumert, Lehmann, Lehrke, Schmitz, Clausen, Hosenfeld, Köller 1997; bspw. Becker und Neumann 2016; Wagner et al. 2008), wird das Interesse damit nicht als relevantes Merkmal zur Beschreibung von Schulartunterschieden beurteilt.

Auch für die Schüler-Lehrer-Beziehung finden sich unter Kontrolle von Antworttendenzen und unter Einbezug der Varianz abweichende Ergebnisse zu anderen Studien. Die Schüler-Lehrer-Beziehung wurde in anderen Vergleichsstudien als besonders niedrig für Gymnasien und Schulen mit mehreren Bildungsgängen beschrieben (bspw. Hertel et al. 2010; Senkbeil et al. 2004). Ähnliche Ergebnisse wurden hier nur ohne Kontrolle von Antworttendenzen gefunden (siehe Tabelle 15). Werden hingegen Antworttendenzen bei der Berechnung der Schulartmittelwerte berücksichtigt, weisen sowohl Klassen an Gymnasien und Schulen mit mehreren Bildungsgängen als auch Klassen an integrierten Gesamtschulen im Mittel ähnliche Werte auf. Lediglich Klassen an Realschulen haben etwas höhere Werte bei diesem Merkmal. Trotzdem sind die Differenzen zwischen Realschulen und den anderen Schularten als gering zu bewerten. Weshalb entgegen der Ergebnisse aus anderen Studien hier unter Kontrolle von Antworttendenzen und unter Berücksichtigung der Varianz des Merkmals nicht von bedeutenden Unterschieden zwischen den Schularten in Bezug auf das Merkmal Schüler-Lehrer-Beziehung ausgegangen werden kann.

7.2.5 Überzeugungen der Lehrenden zu den Zielen im Mathematikunterricht

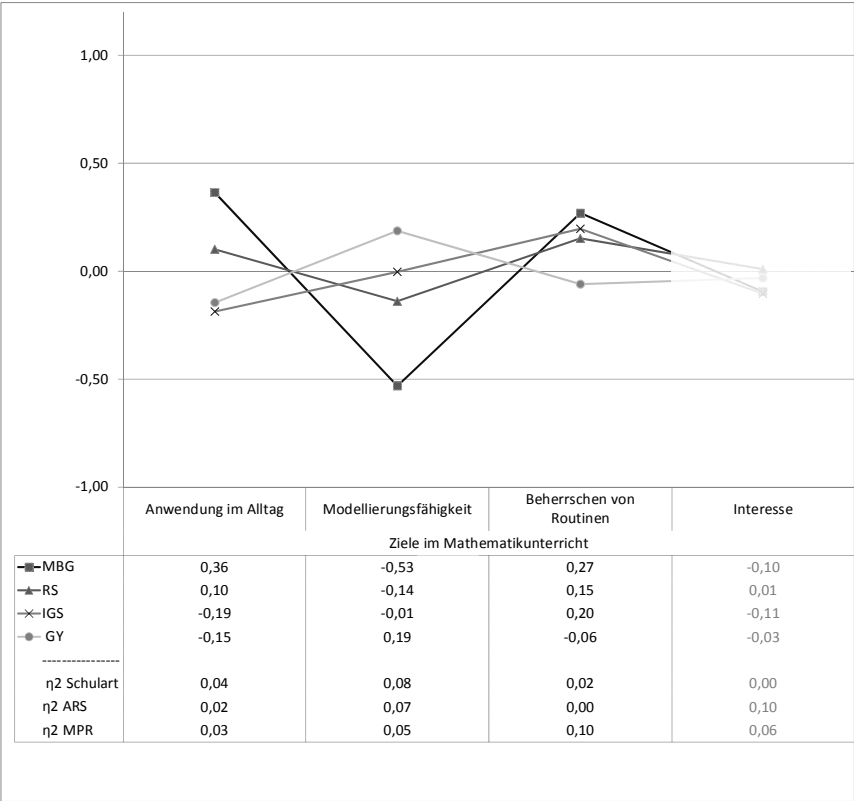


Abbildung 14. Überzeugungen der Lehrenden zu den Zielen des Mathematikunterrichts in den Schularten.

Die Subdimension „Interesse“ ist nicht skalar invariant und daher grau unterlegt. Mittelwertvergleiche sind für diese Subdimension nicht zulässig.
 Abkürzungen: MBG=Schulen mit mehreren Bildungsgängen, RS=Realschulen, IGS=integrierte Gesamtschulen, GY=Gymnasien, ARS=Akquieszenz, MPR=Mid-Point-Responding.

Die Überzeugungen der Lehrenden zu den Zielen des Mathematikunterrichts werden hier nur mit drei der vier Subdimensionen beschrieben. Diese Merkmale werden im Folgenden anhand der Differenzmuster und der Effektstärken

interpretiert. Für die Subdimension „Interesse“ kann nicht davon ausgegangen werden, dass das Konstrukt auf skalarem Niveau vergleichbar zwischen den Schularten gemessen wird. Somit sind Mittelwertvergleiche nicht zulässig (siehe Abschnitt 7.1.1). Diese Subdimension ist in Abbildung 14 aufgeführt und grau unterlegt. Sie wird im Folgenden jedoch nicht vergleichend zwischen den Schularten interpretiert.

Hinsichtlich des Ziels, Mathematik im Alltag anwenden zu können, zeigen sich deutliche Differenzen in den mittleren Überzeugungen der Lehrenden. Dieses Ziel wird besonders von Lehrenden an Schulen mit mehreren Bildungsgängen als wichtig beurteilt. Auch für Lehrende an Realschulen ist die Anwendung im Alltag im Mittel ein wichtiges Ziel. Von geringerer Relevanz ist dieses Ziel im Mittel für Lehrende an Gymnasien und an integrierten Gesamtschulen. Wird bei der Einschätzung der Differenzen zwischen den Schularten jedoch die Varianz in den Überzeugungen einbezogen, zeigt sich mit $\eta^2=0,04$ nur ein geringer Effekt der Schulart. Die Mittelwertdifferenzen sind damit als gering zu beurteilen. Es ist anzunehmen, dass die Überzeugungen bei dieser Dimension innerhalb der Schularten zwischen den Lehrenden stark variieren.

Für das Ziel, Mathematisierungsfähigkeiten zu vermitteln, sind die Differenzen in den mittleren Überzeugungen der Lehrenden zwischen den Schularten ebenfalls deutlich. Die höchste Relevanz schreiben diesem Ziel im Mittel Lehrende an Gymnasien zu. Eine vergleichsweise hohe Relevanz dieses Ziels geben im Mittel auch Lehrende an integrierten Gesamtschulen an. Für Lehrende an Realschulen ist dieses Ziel im Vergleich mit den anderen Schularten im Durchschnitt eher von mittlerer Relevanz. Die geringste Relevanz schreiben diesem Merkmal im Mittel Lehrende an Schulen mit mehreren Bildungsgängen zu. Auch unter Einbezug des Effektstärkemaß mit $\eta^2=0,08$ kann von einem mittleren Effekt der Schulart und somit von bedeutenden Unterschieden in den Überzeugungen der Lehrenden zwischen den Schularten bezüglich der Relevanz von Mathematisierungsfähigkeiten ausgegangen werden.

Für das Ziel, mathematische Routinen und Standardaufgaben zu beherrschen, liegen deutlich geringere Mittelwertdifferenzen vor als bei den anderen Überzeugungen zu den Zielen des Mathematikunterrichts. Lehrende an Schulen mit mehreren Bildungsgängen messen diesem Ziel im Mittel die höchste Relevanz bei. Eine leicht geringere Relevanz hat dieses Ziel im Mittel für Lehrende an integrierten Gesamtschulen und Realschulen. Mit vergleichbar größtem Ab-

stand wird dieses Ziel im Mittel von Lehrenden an Gymnasien als vergleichsweise weniger relevant beurteilt. Die Differenzen zwischen den Schularten sind mit einem marginalen Effekt der Schulart von $\eta^2=0,02$ jedoch als unbedeutend zu beurteilen. Vermutlich stimmen die Lehrenden der verschiedenen Schularten wenig in ihren Überzeugungen bezüglich des Ziels, Routinen und Standardaufgaben zu beherrschen, überein.

Die hier beschriebenen Differenzen in den mittleren Überzeugungen der Lehrenden an den verschiedenen Schularten stimmen auch unter Kontrolle von Antworttendenzen mit den Befunden aus anderen Studien überein. Diese finden für Gymnasien einen Schwerpunkt in kognitiv anspruchsvollen Aufgaben, bei anderen Schularten – besonders an Schulen mit mehreren Bildungsgängen und Hauptschulen – hingegen einen stärkeren Anwendungsbezug und das Einüben von Routinen (Baumert et al. 2004; Helmke, Helmke et al. 2008; bspw. Kunter 2005). Die weitere Analyse (Abschnitt 7.1.2) hatte auch ergeben, dass die Differenzmuster bei diesen Merkmalen nicht durch Antworttendenzen verzerrt werden. Allerdings sind die Differenzen zwischen den Schularten unter Einbezug des Effektstärkemaßes η^2 für zwei der drei Merkmale – „Anwendung im Alltag“ und „Beherrschen von Routinen“ – als marginal einzuschätzen. Die Differenzen zwischen den Schularten wurden bezüglich der Überzeugungen von Lehrenden vermutlich in anderen Studien überschätzt. Die Lehrenden scheinen hier über die Schularten hinweg in ihren Überzeugungen bezüglich dieser Ziele unterschiedliche Auffassungen zu haben, die jedoch nicht auf die Schularten zurückgeführt werden können. In den Überzeugungen der Lehrenden zur Relevanz von Modellierungsfähigkeiten ist hingegen übereinstimmend mit anderen Studien (bspw. Kunter 2005) von bedeutenden Unterschieden zwischen den Schularten auszugehen. Hier wird der hervorstechende Anspruch der Gymnasien in der kognitiven Förderung der Lernenden übereinstimmend mit anderen Studien (bspw. Baumert et al. 2004; Baumert, Lehmann, Lehrke, Schmitz, Clausen, Hosenfeld, Köller 1997) deutlich.

7.2.6 Inhalte des Mathematikunterrichts

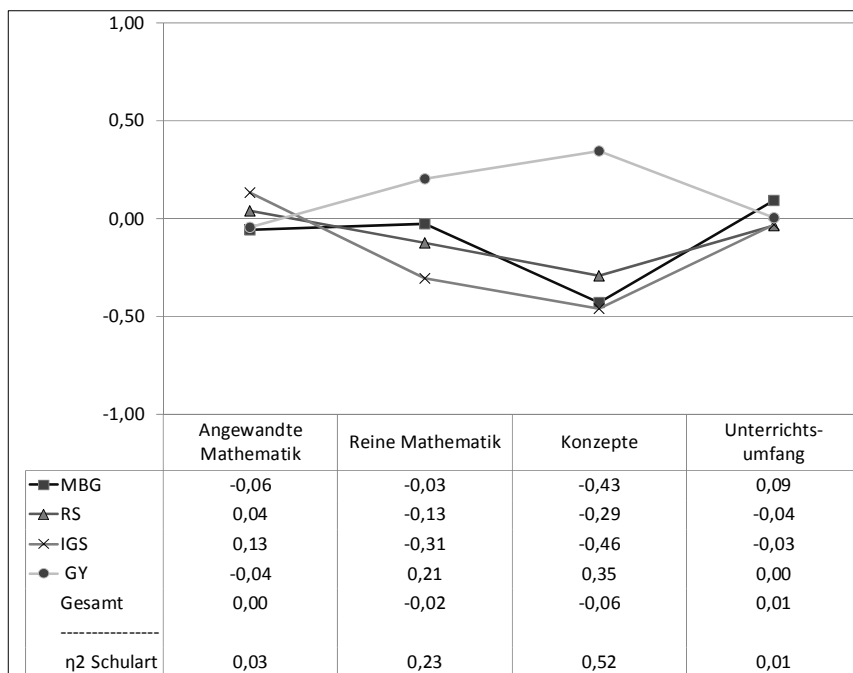


Abbildung 15. Inhalte des Mathematikunterrichts in den Schularten.

Abkürzungen: MBG=Schulen mit mehreren Bildungsgängen, RS=Realschulen, IGS=integrierte Gesamtschulen, GY=Gymnasien, ARS=Akquieszenz, MPR=Mid-Point-Responding.

Die Inhalte des Mathematikunterrichts werden hier über die Angaben von Lernenden zu ihrem Mathematikunterricht beschrieben. Die Lernenden haben zum einen die Häufigkeit von anwendungsbezogenen Mathematikaufgaben (bspw. den wöchentlichen Verbrauch eines Elektrogerätes berechnen) zum anderen die Häufigkeit von reinen Mathematikaufgaben (bspw. eine Gleichung lösen wie $6x + 5 = 29$) eingeschätzt. Die Lernenden wurden ebenfalls gefragt, wie vertraut sie mit verschiedenen mathematischen Konzepten sind (bspw. mit Exponentialfunktionen), was hier als Indikator für die Vermittlung von Fachsprache und Fachkonzepten im Unterricht herangezogen wird. Zudem geben die Lernenden an, wieviel Mathematikunterricht sie pro Woche haben.

Die Unterrichtsstunden der Klassen verschiedener Schularten unterscheiden sich im Mittel nur gering darin, wie häufig anwendungsbezogene Aufgaben bearbeitet werden. Am häufigsten werden solche Aufgaben an integrierten Gesamtschulen eingesetzt. Auch in Realschulen sind diese Aufgaben etwas häufiger präsent als an Schulen mit mehreren Bildungsgängen und Gymnasien. Die Differenzen zwischen den Schularten sind jedoch mit einem Effekt der Schulart von $\eta^2=0,03$ als marginal zu betrachten. Im Mathematikunterricht scheinen anwendungsbezogene Aufgaben in relativ ähnlichem Umfang in allen Schularten eingesetzt zu werden. Dieses Ergebnis stimmt mit den Überzeugungen der Lehrenden überein, welche sich in dem Ziel, die Anwendung im Alltag zu vermitteln, zwischen den Schularten kaum unterscheiden (siehe Abschnitt 3.2.1).

Größere Unterschiede gibt es hingegen bei rein mathematischen Aufgaben. Diese werden in Klassen an Gymnasien deutlich häufiger bearbeitet als an den anderen Schularten. Bei Klassen in Schulen mit mehreren Bildungsgängen und Realschulen werden im Durchschnitt reine Mathematikaufgaben im Vergleich mit mittlerer Häufigkeit durchgenommen. Vergleichsweise selten sind reine Mathematikaufgaben Teil des Unterrichts an integrierten Gesamtschulen. Der hohe Effekt von $\eta^2=0,23$ unterstreicht die große Bedeutung der Schulart für die Anwendung von reinen Mathematikaufgaben im Unterricht. Hier wird der kognitive Anspruch der Gymnasien deutlich, wie er sich auch bei den Zielen der Lehrenden zum Mathematikunterricht gezeigt hat (siehe Abschnitt 3.3.3).

Bei den Gesamtschulen stimmen die Überzeugungen der Lehrenden und die implementierten Unterrichtsinhalte jedoch nicht ganz überein. Lehrende an integrierten Gesamtschulen hatten vergleichsweise hohe Ansprüche, Modellierungsfähigkeiten zu vermitteln. Die Lernenden in Klassen an integrierten Gesamtschulen berichten hingegen im Mittel die niedrigsten Häufigkeiten reiner Mathematikaufgaben im Unterricht. Lernende in Klassen an Schulen mit mehreren Bildungsgängen berichten hingegen im Mittel einen vergleichsweise hohen Anteil an reinen Mathematikaufgaben, obwohl die Lehrenden das Ziel, Modellierungsfähigkeiten zu vermitteln, im Mittel gering bewertet haben. Allerdings ist das Lösen von Gleichungen auch nicht mit Modellierungsfähigkeiten gleichzusetzen.

Am deutlichsten zeigen sich Unterschiede im Mathematikunterricht zwischen den Schularten in der Vermittlung von Fachsprache und Fachkonzepten. Fach-

sprache und Fachkonzepte haben im Vergleich zu den anderen Schularten an Gymnasien im Mittel einen besonders hohen Stellenwert. Deutlich geringere Bedeutung hat die Vermittlung von Fachsprache und Fachkonzepten an Klassen in Realschulen. Die geringste Bedeutung kommt der Fachsprache und den Fachkonzepten an integrierten Gesamtschulen und Schulen mit mehreren Bildungsgängen zu. Der Effekt der Schulart ist mit $\eta^2=0,52$ sehr hoch. In Klassen an den verschiedenen Schularten unterscheidet sich der Unterricht somit stark bezüglich der Vermittlung von Fachkonzepten und Fachsprache, wobei besonders die Gymnasien, welche einen starken Fokus auf die wissenschaftspropädeutische Bildung haben, von den anderen Schularten abweichen.

Erwartungsgemäß gering unterscheidet sich der Unterrichtsumfang zwischen den Schularten. Hier sind bereits durch die sehr ähnlichen länderspezifischen Vorgaben für die Schularten keine Unterschiede zu erwarten. Im realisierten Curriculum, also dem tatsächlichen Unterricht, ist der Zeitumfang in Klassen an Schulen mit mehreren Bildungsgängen geringfügig höher als an den anderen Schularten. Jedoch scheinen die Differenzen mit einem Effekt der Schulart von $\eta^2=0,01$ eher zufällig als systematisch zu sein. Es ist zu vermuten, dass die Varianz im Unterrichtsumfang eher auf Ebene der Klassen oder Schulen zu finden ist, bspw. durch bestimmte Förderangebote der Schulen oder Personal-mangel. Es verwundert daher nicht, dass die eher zufällig anmutenden Differenzen zwischen den Schularten nicht mit den Ergebnissen von Drechsel und Senkbeil (2004) übereinstimmen, welche im Mittel für Gymnasien den geringsten Unterrichtsumfang mit PISA-2003-Daten finden.

Bezüglich der Inhalte des Mathematikunterrichts stimmen die Differenzmuster der Schularten mit den theoretischen Annahmen und den Befunden aus anderen Studien überein (bspw. Neubrand et al. 2011). Der Unterricht an Gymnasien ist durch den hohen kognitiven Anspruch und das wissenschaftspropädeutische Bildungsziel im Vergleich zu den anderen Schularten besonders durch rein mathematische Aufgaben und die Vermittlung von Fachsprache und Fachkonzepten geprägt. Entgegen der Annahmen in anderen Studien scheint sich der Anwendungsbezug jedoch nicht systematisch zwischen den Schularten zu unterscheiden. Zwar finden sich Differenzen zwischen den Schularten in erwarteter Weise, diese sind jedoch nur als marginal zu bewerten. Ein Anwendungsbezug ist somit unabhängig von der Schulart Teil des Mathematikunterrichts der 9. Klassen. Zusätzliche Schwerpunkte in reiner Mathematik und in der Vermitt-

lung von Konzepten und Fachsprachen findet jedoch überwiegend an Gymnasien statt.

7.2.7 Unterrichtsprozesse im Fach Mathematik

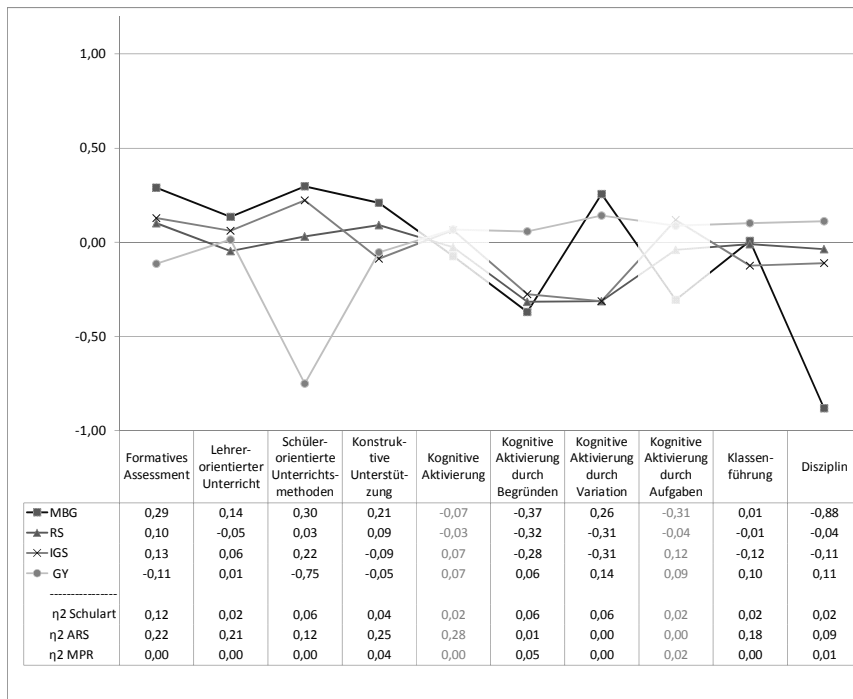


Abbildung 16. Unterrichtsprozesse im Fach Mathematik in den Schularten.

Konstrukte, welche nicht skalar invariant sind, sind grau unterlegt. Mittelwertvergleiche sind für diese Konstrukte nicht zulässig.

Abkürzungen: MBG=Schulen mit mehreren Bildungsgängen, RS=Realschulen, IGS=integrierte Gesamtschulen, GY=Gymnasien, ARS=Akquieszenz, MPR=Mid-Point-Responding.

Acht Merkmale werden in dieser Studie herangezogen, um Unterrichtsprozesse im Mathematikunterricht der 9. Klassen zu beschreiben. Diese Merkmale werden im Folgenden anhand der Differenzmuster und der Effektstärken interpretiert. Bezüge zu Ergebnissen aus anderen Studien, wie sie in Abschnitt 3.5.4 zusammengefasst sind, werden im Anschluss hergestellt. Für zwei Skalen hat die Invarianzanalyse ergeben, dass Mittelwertvergleiche nicht zulässig sind

(siehe Abschnitt 7.1.1.1). Diese zwei Merkmale sind in Abbildung 16 aufgeführt und grau unterlegt. Sie werden im Folgenden jedoch nicht interpretiert.

Formatives Assessment findet im Mittel am häufigsten in Klassen an Schulen mit mehreren Bildungsgängen statt. Lernende in Klassen an dieser Schulart bekommen demnach aus ihrer Sicht vergleichsweise häufig Rückmeldung über ihren Lernstand und ihre Entwicklungsmöglichkeiten. Im Unterricht an integrierten Gesamtschulen und Realschulen wird diese Methode im Durchschnitt vergleichsweise mit mittlerer Häufigkeit eingesetzt. Die geringste Rückmeldung zum Lernstand und den Entwicklungsmöglichkeiten erhalten Lernende in Klassen an Gymnasien. Hier scheint von den Lernenden mehr Selbstreflexion erwartet zu werden. Mit einem mittleren Schularteffekt von $\eta^2=0,12$ sind diese Unterschiede zwischen den Schularten als bedeutend zu beurteilen.

Der Unterricht der Lernenden unterscheidet sich in der mittleren Ausprägung der *Lehrerorientierung* entgegen der Erwartungen nur gering zwischen den Schularten. Die Differenzmuster zeigen zwar, dass sich der Unterricht an Schulen mit mehreren Bildungsgängen und integrierten Gesamtschulen im Mittel durch eine leicht stärkere Steuerung durch die Lehrkraft auszeichnet als in Klassen an Gymnasien und Realschulen. Die Differenzen sind jedoch vergleichsweise gering zwischen den Schularten. Auch das Effektstärkemaß der Schulart $\eta^2=0,02$ deutet an, dass die Unterschiede zwischen den Schularten marginal sind.

Schülerorientierte Unterrichtsmethoden werden im Mittel am häufigsten in Klassen an Schulen mit mehreren Bildungsgängen und integrierten Gesamtschulen eingesetzt. Im Unterricht in Klassen an diesen Schularten findet vergleichsweise häufig Projektarbeit und Kleingruppenarbeit statt. Zudem werden Lernende vergleichsweise häufig in die Unterrichtsplanung einbezogen und Aufgaben werden vergleichsweise häufig differenziert entsprechend der Fähigkeiten der Lernenden verteilt. Im Unterricht an Realschulen finden schülerorientierte Unterrichtsmethoden im Durchschnitt vergleichsweise mit mittlerer Häufigkeit statt. In Klassen an Gymnasien werden schülerorientierte Unterrichtsmethoden am seltensten eingesetzt. Mit einem mittleren Schularteffekt von $\eta^2=0,16$ sind diese Unterschiede zwischen den Schularten als bedeutend zu beurteilen.

Auch in der *konstruktiven Unterstützung* durch die Lehrperson unterscheiden sich die Klassen der verschiedenen Schularten in den mittleren Ausprägungen kaum voneinander. Besonders in Klassen an Gymnasien und integrierten Gesamtschulen ist die konstruktive Unterstützung durch die Lehrperson im Mittel ähnlich gering. Obwohl Lernende in Klassen an Realschulen und Schulen mit mehreren Bildungsgängen geringfügig höhere mittlere Ausprägungen in der konstruktiven Unterstützung berichten, ist mit einem Effekt der Schulart von $\eta^2=0,04$ nicht von bedeutenden Unterschieden zwischen den Schularten in der konstruktiven Unterstützung durch den Lehrenden auszugehen.

Interessante Unterschiede im Differenzmuster zeigen sich für die *kognitive Aktivierung durch Begründen*. Hier heben sich die Klassen an Gymnasien im Mittel deutlich von den anderen Schularten ab. Lehrende an Gymnasien geben somit im Mittel deutlich häufiger an, dass sie die Lernenden dazu anhalten, Aussagen zu begründen. Die niedrigsten Ausprägungen liegen bei diesem Merkmal im Mittel für Schulen mit mehreren Bildungsgängen vor. Aber auch Realschulen und integrierte Gesamtschulen haben ähnlich niedrige Werte. Mit einem mittleren Schularteffekt von $\eta^2=0,6$ sind diese Unterschiede zwischen den Schularten als bedeutend zu beurteilen, wobei hier vermutlich vor allem die Unterschiede zwischen Lehrenden an Gymnasien im Vergleich zu den anderen Schularten ausschlaggebend sind, weniger die – kaum vorhandenen – Differenzen zwischen Lehrenden an Realschulen, integrierten Gesamtschulen und Schulen mit mehreren Bildungsgängen.

Anders verhält es sich bei der *kognitiven Aktivierung durch Variation* der Aufgaben. Hier geben die Lehrenden an Schulen mit mehreren Bildungsgängen und an Gymnasium im Mittel vergleichsweise häufig an, die Aufgabenstellung bzw. Einkleidung der Aufgaben zu variieren. Für Lehrende an integrierten Gesamtschulen und Realschulen ist dies im Mittel seltener eine Methode zur kognitiven Aktivierung ihrer Klasse. Mit einem mittleren Effekt der Schulart von $\eta^2=0,06$ sind die Unterschiede zwischen Gymnasien und Schulen mit mehreren Bildungsgängen einerseits und Realschulen und integrierten Gesamtschulen andererseits als bedeutend einzuschätzen.

Die *Klassenführung* scheint zwischen den Klassen der Schularten im Mittel relativ ähnlich zu sein. Lernende an Gymnasien geben im Mittel geringfügig häufiger an, dass ihre Mathematiklehrkraft die Unterrichtsstunde strukturiert führt. Klassen an Realschulen und Schulen mit mehreren Bildungsgängen lie-

gen bei diesem Merkmal im Mittelfeld. Klassen an integrierten Gesamtschulen haben im Mittel die geringsten Werte. Die Unterschiede sind jedoch mit einem Effekt der Schulart von $\eta^2=0,02$ als marginal zu beurteilen. Innerhalb der Schularten scheint sich die Beurteilung der Klassenführung zwischen den Klassen somit in relevanterem Ausmaß zu unterscheiden, so dass keine systematischen Unterschiede zwischen den Schularten angenommen werden können.

In engem Bezug zur Klassenführung steht die *Disziplin* der Klasse, teilweise wird die Disziplin als Teilaspekt der Klassenführung gesehen (bspw. Lipowsky 2015). Entsprechend zeigen sich für die Gymnasien, die Realschulen und die integrierten Gesamtschulen ähnliche mittlerer Ausprägungen in diesem Merkmal wie bei der Klassenführung. Bedeutend weichen jedoch die mittleren Urteile der Lernenden in Schulen mit mehreren Bildungsgängen von anderen Schularten ab. Trotz der starken mittleren Differenz zu den anderen Schularten kann dieser Unterschied nicht als bedeutend eingeschätzt werden. Der Effekt der Schulart von $\eta^2=0,02$ gibt starke Hinweise darauf, dass die Varianz der Klassen innerhalb der Schularten sehr hoch ist und somit keine systematischen Unterschiede zwischen den Schularten angenommen werden können.

Die hier beschriebenen Mittelwerte von Unterrichtsprozessmerkmalen wurden unter Kontrolle von Antworttendenzen berechnet und schließen bei der Interpretation das Effektstärkemaß η^2 ein. Es interessiert daher, inwieweit die Ergebnisse dieser Studie von den Ergebnissen in anderen Studien abweichen und ob dies auf die abweichende Berechnung und Interpretation zurückgeführt werden kann. Deshalb werden im Folgenden die beschriebenen Differenzmuster und die Bedeutung der Schulart in Bezug zum Forschungsstand, wie er in Abschnitt 3.5.4 beschrieben wurde, interpretiert. Der Vergleich ist jedoch eingeschränkt, da die Merkmale, welche Unterrichtsprozesse beschreiben, häufig nur vergleichend zwischen Gymnasien und Hauptschulen beschrieben wurden, oder Gymnasien den anderen Schularten gemeinsam gegenüber gestellt werden. Somit sind differenzierte Ergebnisse für die einzelnen Schularten selten. Zudem werden Schulen mit mehreren Bildungsgängen häufig nicht als eigenständige Schulart erfasst, sondern die drei Bildungsgänge den Schularten Hauptschule, Realschule und Gymnasium zugeordnet. Dadurch fehlen häufig Vergleichskriterien für diese Schulart.

Für fünf Merkmale werden auch unter Kontrolle von Antworttendenzen keine Abweichungen in den Differenzmustern gefunden: Klassenführung und Diszi-

plin, kognitive Aktivierung durch Begründen sowie formatives Assessment und schülerorientierte Unterrichtsmethoden. Für die Klassenführung ändert sich zwar unter Kontrolle von Antworttendenzen das Differenzmuster (siehe Abschnitt 7.1.2), trotzdem weichen die hier gefundenen Ergebnisse mit und ohne Kontrolle von Antworttendenzen nicht von Befunden aus anderen Studien ab (Bayer et al. 2015; bspw. Gruehn 2000; Klieme und Rakoczy 2001; Kunter et al. 2005), da differenzierte Beschreibungen der Schularten fehlen. Den Erkenntnissen anderer Studien, dass an Gymnasien eine vergleichsweise starke Struktur vorherrscht (durch effektive Nutzung der Zeit, klar definierte Regeln, stabile Routinen und Unterrichtsdisziplin), an Hauptschulen hingegen die niedrigsten Werte vorliegen, können hier Befunde für die anderen Schularten ergänzt werden: Klassen in Schulen mit mehreren Bildungsgängen und in Realschulen haben im Mittel sehr ähnliche Ausprägungen beim Merkmal Klassenführung. Für beide Schularten sind die mittleren Ausprägungen niedriger als am Gymnasium. Die niedrigsten Ausprägungen liegen im Mittel für Klassen an integrierten Gesamtschulen vor. Allerdings wurden die Unterschiede zwischen den Schularten bislang vermutlich überschätzt, wie sich unter Einbezug des Effektstärkemaßes η^2 zeigt. Die Varianz der Klassen innerhalb der Schularten scheint sehr hoch zu sein, weshalb Unterschiede zwischen Klassen und Schulen vermutlich bedeutsamer sind, als Unterschiede zwischen den Schularten und somit nicht davon auszugehen ist, dass sich die Klassenführung zwischen den Schularten bedeutsam unterscheidet.

Nahezu analog ist die Ergebnislage beim Merkmal Disziplin, welches auch als Teilaspekt der Klassenführung gesehen werden kann (bspw. Lipowsky 2015). In Anbetracht der Unschärfe bei der Beschreibung in anderen Studien (bspw. Tillmann und Meier 2001) ist das Differenzmuster in der hier vorliegenden Studie als übereinstimmend mit der Befundlage in anderen Studien zu beurteilen. Mit Kontrolle von Antworttendenzen werden die Differenzen in den mittleren Ausprägungen zwischen den Schularten jedoch deutlich größer, vor allem der vergleichsweise stark negative Wert bei den integrierten Gesamtschulen sticht hervor. Trotz deutlicher Unterschiede in den mittleren Ausprägungen beim Merkmal Disziplin ist jedoch nicht von bedeutenden Unterschieden zwischen den Schularten auszugehen. Vielmehr ist zu vermuten, dass Unterschiede zwischen Klassen und Schulen bedeutender sind. Somit ist anzunehmen, dass Unterschiede in der Disziplin zwischen den Schularten bislang überschätzt wurden. Trotz der großen Differenzen in den mittleren Ausprägungen ist davon auszugehen, dass die Unterschiede zwischen den Klassen innerhalb der Schul-

arten groß sind und somit nicht von bedeutenden Unterschieden zwischen den Schularten ausgegangen werden sollte.

Das Merkmal formatives Assessment wurde anscheinend bislang nicht vergleichend zwischen den Schularten ausgewertet. Als eine Methode, welche eingesetzt werden kann, um die Lernenden in ihrer Entwicklung zu unterstützen und Unterricht an die Bedürfnisse von Lernenden anzupassen, werden daher die Erkenntnisse zur konstruktiven Unterstützung von Lernenden sowie zu schülerorientierten Unterrichtsmethoden als Vergleichsgrundlage herangezogen. Aus anderen Studien ist hier bekannt, dass Hauptschulen in diesen Merkmalen im Mittel die höchsten Ausprägungen haben, Gymnasien die niedrigsten (bspw. Gruehn 2000; Klieme, Schümer et al. 2001; Kunter 2005). Die anderen Schularten liegen im mittleren Bereich. In dieser Studie liegt für das formative Assessment ein analoges Befundmuster vor. Die Unterschiede zwischen den Schularten sind bezogen auf das Effektstärkemaß η^2 bedeutend, weshalb es erstaunt, dass das formative Assessment bislang anscheinend nicht in Schulartvergleichen berücksichtigt wurde.

Besonders die kognitive Aktivierung wird in anderen Studien vorwiegend zwischen Gymnasien und Hauptschulen kontrastiert, was den Vergleich zwischen den Befundmustern zu der hier vorliegenden Studie erschwert. Als Kriterium wird daher die dominierende Position der Gymnasien herangezogen. Gymnasien wird ein besonders hohes Ausmaß an kognitiv aktivierendem Unterricht zugeschrieben. Das hier gefundene Differenzmuster zur kognitiven Aktivierung durch Begründen fügt sich perfekt in diese Befundlage ein. Mit Abstand hat dieses Merkmal bei den Gymnasien die höchsten Ausprägungen. Der mittlere Effekt der Schulart mit $\eta^2=0,06$ bekräftigt die angenommenen Unterschiede zwischen den Gymnasien und den anderen Schularten in der kognitiven Aktivierung durch Begründen.

Unterschiede im Differenzmuster zeigt sich jedoch für das Merkmal kognitive Aktivierung durch Variation. Hier haben Gymnasien im Mittel zwar höhere Ausprägungen als Realschulen und integrierte Gesamtschulen. Allerdings haben Schulen mit mehreren Bildungsgängen in dieser Studie im Mittel die höchsten Werte und somit höhere Werte bei der kognitiven Aktivierung als Gymnasien. Schulen mit mehreren Bildungsgängen wurden in anderen Studien bei der Beschreibung von kognitiver Aktivierung meist nicht separat betrachtet. Deshalb wird hier das gefundene Differenzmuster in dieser Studie als abwei-

chend bewertet. Antworttendenzen sind dabei nicht ausschlaggebend (siehe Abschnitt 7.1.2). Vielmehr könnten tatsächlich inhaltliche Gründe vorliegen: Unter der Annahme, dass Hauptschulen schwerpunktmäßig Grundlagen vermitteln, ist gerade für das Wiederholen und Einüben der Grundlagen eine variierende Einkleidung sich wiederholender Aufgabeninhalte bedeutsam.

Starke Abweichungen im hier gefundenen Differenzmuster zu anderen Forschungsbefunden liegen für die Lehrerorientierung vor. Hier fand Kunter (2005) anhand von TIMSS-Daten, dass lehrerorientierter Unterricht besonders häufig an Gymnasien, am seltensten an integrierten Gesamtschulen und Hauptschulen zu beobachten war. Realschulen nehmen dort eine Mittelposition ein (Kunter 2005, S. 227–228). Für die PISA-Daten zeigen sich in dem oben beschriebenen Differenzmustern zum lehrerorientierten Unterricht jedoch andere Rangfolgen der Schularten. Mit und ohne Kontrolle von Antworttendenzen wird die Lehrerorientierung in Klassen an Schulen mit mehreren Bildungsgängen und integrierten Gesamtschulen im Mittel am höchsten beurteilt. Für Gymnasien und Realschulen liegen im Mittel die niedrigsten Einschätzungen vor. Die Unterschiede in den Ergebnissen sind vermutlich inhaltlich begründet: In der Studie von Kunter (2005) wird die Lehrerorientierung beschrieben über die Störungsfreiheit und die Anregungen von Lehrprozessen, bei denen an Vorwissen und Schülervorstellungen angeknüpft wird (Kunter 2005, S. 214). Letzteres kann auch als eine Form der kognitiven Aktivierung betrachtet werden. In der eingesetzten PISA-Skala liegt der Schwerpunkt lehrerorientierten Unterrichts hingegen im fragend geführten Unterrichtsgespräch und der Kommunikation von Zielvorgaben des Lehrenden an die Lernenden (OECD 2014, S. 330). Unterricht in diesem Sinne scheint unabhängig von der Schulart typisch für den Unterricht an den Sekundarschulen in Deutschland zu sein (Klieme, Schümer et al. 2001).

Abweichungen in den Befunden finden sich auch für das Merkmal konstruktive Unterstützung. Hier wird in anderen Forschungsarbeiten übereinstimmend berichtet, dass Hauptschulen im Mittel die höchsten Ausprägungen bei diesem Merkmal aufweisen, Gymnasien die niedrigsten. Die anderen Schularten befinden sich, soweit analysiert, im Mittelfeld (Bayer et al. 2015; bspw. Gruehn 2000; Klieme und Rakoczy 2001). Ohne Kontrolle von Antworttendenzen findet sich dieses Befundmuster auch in der hier vorliegenden Arbeit (siehe Abschnitt 7.1.2). Werden hingegen Antworttendenzen kontrolliert, haben Klassen an integrierten Gesamtschulen im Mittel leicht niedrigere Werte in der

wahrgenommenen konstruktiven Unterstützung als Klassen an Gymnasien. Die höchsten Werte berichten im Mittel Lernende in Klassen an Schulen mit mehreren Bildungsgängen. Realschulen liegen im Mittelfeld. Die vergleichsweise geringen Werte in der eingeschätzten konstruktiven Unterstützung für die integrierten Gesamtschulen erstaunen, da gerade dieser Schulform aus der Entwicklungsgeschichte und dem Selbstverständnis heraus eine besonders starke Orientierung am Individuum und individuelle Förderung zugeschrieben wird (siehe Kapitel 2). Lehrerhandeln, das sensitiv für die individuellen Bedürfnisse der Lernenden ist und unterstützt, sollte daher gerade an dieser Schulform stark ausgeprägt sein. Eventuell sind es aber gerade solch hohe Erwartungen der Lernenden an ihre Lehrkräfte, die zu dieser sehr gering wahrgenommenen konstruktiven Unterstützung führen. Hier besteht weiterer Forschungsbedarf. Zum einen bedarf es Studien, welche die wahrgenommene Unterstützung unter Kontrolle von Antworttendenzen untersuchen. Zum anderen würden qualitative Erkenntnisse, bspw. aus Unterrichtsbeobachtungen, das Verständnis von der wahrgenommenen konstruktiven Unterstützung vergleichend zwischen Schularten erhöhen. Bei der Interpretation ist jedoch Vorsicht geboten, um die Unterschiede zwischen den Schularten nicht zu überschätzen. Der Effekt der Schulart ist mit einem $\eta^2=0,04$ noch als gering zu beurteilen. Weshalb vermutlich die Unterschiede innerhalb der Schularten zwischen Klassen relativ hoch sind.

Zusammenfassend liegen in dieser Arbeit für vier Merkmale (Klassenführung und Disziplin, kognitive Aktivierung durch Begründen und formatives Assessment) Differenzmuster vor, wie sie auch in anderen Arbeiten gefunden wurden. Für die Merkmale kognitive Aktivierung durch Variation, Lehrerorientierung und konstruktive Unterstützung weichen die Differenzmuster hingegen von den Ergebnissen anderer Studien ab. Die Gründe hierfür sind sehr unterschiedlich. Für die kognitive Aktivierung durch Variation sind vermutlich Unschärfen bei der Beschreibung in anderen Studien ausschlaggebend. Für die Lehrerorientierung führen vermutlich abweichende Definitionen des Merkmals zu den konträren Ergebnissen. Bei der Schülerorientierung sind Antworttendenzen, welche vermutlich in anderen Studien nicht berücksichtigt wurden, die Ursache abweichender Differenzmuster zu den Ergebnissen in der hier vorliegenden Arbeit. Bei fünf von sieben Merkmalen, welche Unterrichtsprozesse beschreiben, ist jedoch nicht davon auszugehen, dass die Mittelwertunterschiede bedeutsam sind zwischen den Schularten. Daher wurden die Unterschiede zwischen den Schularten bei diesen Merkmalen vermutlich bislang überschätzt. Bedeutende

Mittelwertunterschiede zwischen den Schularten zeigen sich hier nur für die Merkmale kognitive Aktivierung durch Begründen und formatives Assessment.

7.2.8 Zusammenfassende Interpretation

Die deskriptiven Beschreibungen der Schularten zeigen einmal mehr, wie stark sich die Lernenden zwischen den Schularten in ihren Leistungen, der Leistungsentwicklung und den sozioökonomischen Hintergrundmerkmalen unterscheiden. Die konträren Ergebnisse der beiden Leistungstests weisen jedoch darauf hin, dass die Lernenden der verschiedenen Schularten von der neunten zur zehnten Klasse Fortschritte in verschiedenen Schwerpunktbereichen der Mathematik erzielen. Lernende in Gymnasien steigern ihre Leistung eher in den in PISA erhobenen Fähigkeiten des reflektierenden und anwendungsbezogenen Denkens. Lernende in den anderen Schularten, besonders an den integrierten Gesamtschulen und Schulen mit mehreren Bildungsgängen, zeigen die größten Leistungssteigerungen in den curriculumbasierten Aufgaben des BiSta-Tests.

Bedeutende Unterschiede zwischen den Schularten finden sich auch bei den Inhalten des Mathematikunterrichts in der Häufigkeit von Aufgaben mit rein mathematischen Inhalten im Unterricht und in der Vermittlung von Fachsprache und Fachkonzepten. Das Gymnasium hat hier im Vergleich zu den anderen Schularten eine Sonderstellung, indem diese Inhalte vergleichsweise häufig Teil des Mathematikunterrichts sind. Entgegen der Erwartungen scheinen anwendungsbezogene Aufgaben in allen Schularten ähnlich häufig vorzukommen.

Nur wenige der Merkmale, welche die Überzeugungen der Lehrenden, affektive Schülermerkmale und Unterrichtsprozesse erfassen, unterscheiden sich bedeutend zwischen den Schularten. Auch wenn sich teils deutliche Differenzen in den mittleren Ausprägungen einzelner Merkmale zwischen den Schularten zeigen, ist der Effekt der Schulart bei dem Großteil der Merkmale gering. Unklar bleibt, ob die verwendeten Skalen nicht geeignet sind, die Unterschiede zwischen den Schularten zu erfassen, oder ob, rekurrierend auf Fend, für diese Merkmale der Einzelschule eine größere Bedeutung zukommt. Letzteres ist, im Rückgriff auf die Schulkulturtheorie Helspers (2001), besonders naheliegend für die affektiven Schülermerkmale.

Die Merkmale, welche sich in dieser Studie als bedeutend für den Vergleich und die Beschreibung von Unterricht erwiesen haben, sind in Abbildung 17 zusammengefasst. Die Abbildung verdeutlicht die Sonderstellung der Gymnasien. In nahezu allen Merkmalen weichen die Klassenmittelwerte sichtlich von den anderen Schularten ab. Darüber hinaus sind auch Differenzen zwischen den anderen Schularten offensichtlich, wobei die Rangfolge der Schularten merkmalsabhängig variiert. Beispielsweise scheint jede nicht gymnasiale Schulart ihren eigenen Schwerpunkt bezüglich des Mathematikunterrichts zu haben: integrierte Gesamtschulen in dem Ziel, Modellierungsfähigkeit zu vermitteln, Schulen mit mehreren Bildungsgängen im Einsatz rein mathematischer Aufgaben und Realschulen in der Vermittlung von Fachsprache und Fachkonzepten. Es ist daher lohnenswert, Unterschiede zwischen allen Schularten detailliert zu beschreiben. Schularten bei Analysen gar zusammenzufassen, empfiehlt sich aufgrund der teilweise beträchtlichen Unterschiede zwischen den Schularten nicht.

Der Überblick von Merkmalen, welche sich bedeutend zwischen den Schularten unterscheiden, verdeutlicht auch den inhaltlichen Schwerpunkt der Schulartunterschiede. Unterschiede zeigen sich – wie zu erwarten war – vor allem in der Zusammensetzung der Schülergruppe, bzw. der Klassen in Bezug auf die Leistung und sozioökonomische Merkmale. Darüber hinaus nehmen Merkmale, welche die kognitive Aktivierung adressieren, eine bedeutende Rolle ein. So sind Unterschiede zwischen den Schularten in den Überzeugungen der Lehrenden, den Unterrichtsinhalten und den Unterrichtsprozessen bezüglich der kognitiven Aktivierung und der Vermittlung von fachwissenschaftlichen Konzepten deutlich. Durch die großen Mittelwertdifferenzen im formativen Assessment und in schülerorientierten Unterrichtsmethoden sind über die Schülergruppe und die kognitive Aktivierung hinaus auch bedeutende Unterschiede im unterstützenden Umgang mit den Lernenden zu vermuten.

Ausgehend von den Befunden der deskriptiven Analysen interessieren nun Zusammenhänge zwischen den Merkmalen. Eine Hypothese dieser Arbeit ist, dass die Unterschiede in den Unterrichtsprozessen vermittelt werden über Merkmale der Schülergruppe, die Überzeugungen der Lehrenden und die Unterrichtsinhalte. Diese Annahme soll im nächsten Schritt geprüft werden.

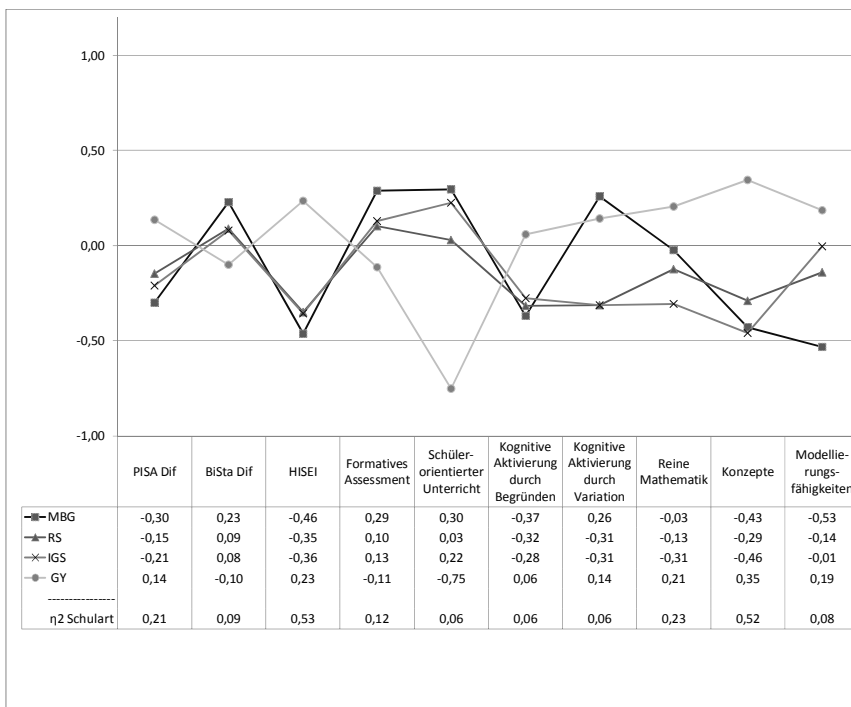


Abbildung 17. Zusammenfassende Darstellung bedeutender Schulartunterschiede.

Auf die Darstellung der absoluten Mathematikleistung und des sozioökonomischen Status wurde an dieser Stelle verzichtet. Die mehrfach replizierten Unterschiede zwischen den Schularten hinsichtlich dieser Merkmale, die sich auch in den hier verwendeten Daten zeigen, werden als bekannt vorausgesetzt und können der Zusammenfassung keine Impulse geben. Auch auf die Darstellung der Antworttendenzen wurde an dieser Stelle verzichtet.

Abkürzungen: MBG=Schulen mit mehreren Bildungsgängen, RS=Realschulen, IGS=integrierte Gesamtschulen, GY=Gymnasien.

7.3 Schulartunterschiede in den Unterrichtsprozessen mediert über die Überzeugungen der Lehrenden, die Unterrichtsinhalte und die Schülergruppe

Unterrichtsprozesse unterscheiden sich zwischen den Schularten. Dies legen die deskriptiven Befunde dieser Arbeit nahe, welche damit an die Ergebnisse aus anderen Untersuchungen zu Schulartunterschieden anknüpfen (bspw. Baumert et al. 2004; Kunter 2005). Bislang offen ist die Frage, wie bzw. über welche Mechanismen die Schularten auf die Unterrichtsprozesse wirken. Ziel der folgenden Analysen ist es, dieser Frage empirisch nachzugehen. Theoretisch wurden in dieser Arbeit bereits vermittelnde Mechanismen auf Seiten der Lehrenden, der Schülergruppe sowie der Unterrichtsinhalte hergeleitet (Kapitel 3) und die Ausprägungen der jeweiligen Merkmale zwischen den Schularten untersucht (Abschnitt 7.2). Methodisch wird die Frage nach den vermittelnden Mechanismen in den folgenden Analysen anhand von Mediationsmodellen untersucht.

Ein Effekt des Mediators liegt vor, wenn der direkte Effekt zwischen der Schulart und dem Unterrichtsprozess durch den Einfluss der Mediatorvariablen unterbrochen bzw. minimiert wird und der indirekte Effekt der Schulart auf das Unterrichtsprozessmerkmal signifikant ist (MacKinnon 2012; ausführlicher beschrieben im Methodenteil). Zudem gilt es, vorab zu prüfen, dass signifikante Effekte zwischen der Schulart und den Unterrichtsprozessmerkmalen vorliegen sowie signifikante Effekte zwischen Schulart und der Mediatorvariablen bestehen. Hierfür werden die deskriptiven Beschreibungen der Schularten herangezogen, welche für vier Unterrichtsprozessmerkmale und potentielle Mediatoren bedeutende Unterschiede zwischen den Schularten identifizieren (Abschnitt 7.2.7).

Tabelle 17 fasst diese Merkmale zusammen. Bei den Unterrichtsprozessmerkmalen handelt es sich zum einen um Merkmale, welche Unterrichtsmethoden beschreiben (schülerorientierte Unterrichtsmethoden und formatives Assessment) und zum anderen um Merkmale, welche das Potential zur kognitiven Aktivierung beschreiben (kognitive Aktivierung durch die Variation von Aufgaben und kognitive Aktivierung durch Insistieren auf Begründen). Potentielle Mediatoren finden sich bei den Lehrenden (Unterrichtsziel: Modellierungsfähigkeit), der Schülergruppe (Leistungszusammensetzung und soziale Zusam-

mensetzung) und den Unterrichtsinhalten (reine Mathematik, mathematische Konzepte).

Tabelle 17. Ana9ysierte Unterrichtsprozessmerkmale und potentielle Mediatoren auf Seiten der Lehrenden, der Unterrichtsinhalte und der Schülergruppe.

Mediatoren		Unterrichtsprozesse
Lehrende	Unterrichtsziel: Modellierungsfähigkeit*	Kognitive Aktivierung durch Begründen*
Inhalte	Reine Mathematik	
	Mathematische Konzepte	Kognitive Aktivierung durch Variation*
Schülergruppe	Leistungskomposition	Schülerorientierte Unterrichtsmethoden*
	Soziale Komposition	Formatives Assessment*

Für Merkmale, welche mit einem Stern (*) gekennzeichnet sind, werden in den Analysen die Antworttendenzen der Lehrenden bzw. der Lernenden kontrolliert.

Für die vier Unterrichtsprozessmerkmale wird in den folgenden Analysen untersucht, inwieweit die Schulartunterschiede über die fünf potentiellen Mediatoren vermittelt werden. Hierfür wird zunächst die vermittelnde Rolle jeweils eines Mediators für jeweils ein Unterrichtsprozessmerkmal untersucht. Das Gymnasium ist die Referenzkategorie des Schulartvergleichs. Analog zur deskriptiven Beschreibung der Schularten wurden auch in den Mediationsmodellen Merkmale, welche mittels Likert-Skalen erhoben wurde, unter Kontrolle der Antworttendenzen der Lehrenden und Lernenden analysiert. Die folgende Ergebnisdarstellung berücksichtigt jedoch vereinfacht nur die interessierenden Effekte auf der Klassenebene.

Tabelle 38 bis Tabelle 41 im Online-Anhang (Abschnitt 12.6) fassen die Ergebnisse für alle 20 Mediationsmodelle (vier Unterrichtsprozesse à fünf Mediatorvariablen) zusammen. Die wichtigsten modellübergreifenden Ergebnisse sind hier kurz beschrieben: Über die untersuchten Modelle hinweg sind die direkten Effekte der Schulart auf die jeweilige Mediatorvariable, analog zu den deskriptiven Befunden, signifikant. Die direkten Effekte der Schulart auf die Unterrichtsprozesse sind hingegen nur in wenigen Modellen für einzelne

Schularten signifikant. Nicht signifikante Zusammenhänge indizieren, dass unter Kontrolle der jeweiligen Mediatorvariablen die Zusammenhänge zwischen der Schulart und den Unterrichtsprozessmerkmalen unterbrochen werden. Signifikante indirekte Effekte der Schulart auf Unterrichtsprozessmerkmale, vermittelt über den jeweiligen Mediator, werden über die 95 % Bootstrap Konfidenzintervalle beurteilt (siehe im Methodenteil Abschnitt 6.8). Hier zeigt sich, dass signifikante indirekte Effekte der Schulart auf Unterrichtsprozessmerkmale nur für zwei Unterrichtsprozessmerkmale vorliegen: Der Effekt der Schulart auf den Einsatz schülerorientierter Unterrichtsmethoden wird mediiert über die Leistungskomposition, das Merkmal mathematische Konzepte und über das Merkmal reine Mathematik. Der Effekt der Schulart auf das Potential zur kognitiven Aktivierung durch die Variation der Aufgaben wird mediiert über das Merkmal mathematische Konzepte. Für diese zwei Unterrichtsprozessmerkmale werden die Ergebnisse im Folgenden näher beschrieben.

7.3.1 Potential zur kognitiven Aktivierung durch Variation der Aufgaben

Der Effekt der Schulart auf die kognitive Aktivierung durch Variation der Aufgaben wird mediiert über das Merkmal mathematische Konzepte, wie die Ergebnisse in Abbildung 18 indizieren. Der Effekt der Schulart auf das Potential zur kognitiven Aktivierung durch Variation liegt nicht mehr vor, wenn die Mediatorvariable ins Modell aufgenommen wird. Die Pfade β_{RS} , β_{IGS} und β_{MBG} sind also nicht mehr signifikant. Die Schulart hängt vielmehr negativ mit dem Merkmal mathematische Konzepte zusammen ($\beta_{RS_M}=-0,61^*$; $\beta_{IGS_M}=-0,70^*$; $\beta_{MBG_M}=-0,74^*$), welches wiederum positiv mit dem Potential zur kognitiven Aktivierung zusammenhängt ($\beta_M=0,69^*$). Dies bedeutet, dass in Realschulen, integrierten Gesamtschulen und Schulen mit mehreren Bildungsgängen das Ausmaß, in dem mathematische Konzepte vermittelt werden, geringer ist als an Gymnasien. Dies wurde bereits ausführlich in den deskriptiven Analysen beschrieben. Je geringer das Ausmaß ist, in dem mathematische Konzepte vermittelt werden, desto geringer ist das Potential zur kognitiven Aktivierung durch Variation der Aufgaben. Dadurch ergibt sich ein negativer indirekter Effekt der Schulart auf das Potential zur kognitiven Aktivierung, welcher über das Merkmal mathematische Konzepte vermittelt wird ($\beta_{RS_IND}=-0,42^*$; $\beta_{IGS_IND}=-0,49^*$; $\beta_{MBG_IND}=-0,52^*$).

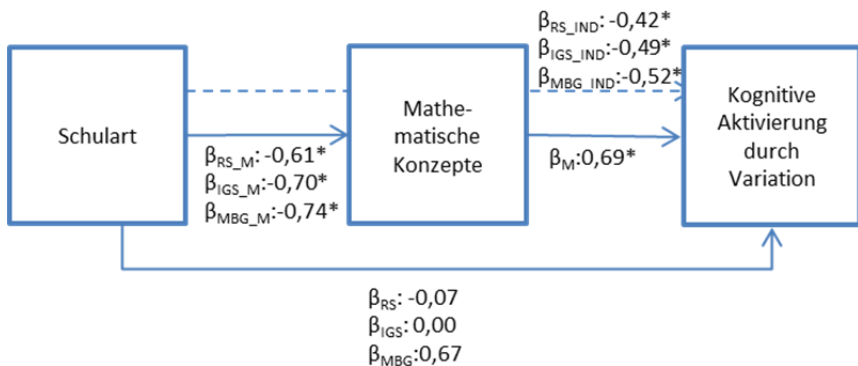


Abbildung 18. Mediation der Schultartunterschiede im Potential zur kognitiven Aktivierung durch Variation der Aufgaben über mathematische Konzepte.

Signifikante Effekte sind mit einem Stern (*) gekennzeichnet.

Abkürzungen: MBG=Schulen mit mehreren Bildungsgängen, RS=Realschulen, IGS=integrierte Gesamtschulen, M=Mediator, IND=indirekter Effekt ---->.

7.3.2 Schülerorientierte Unterrichtsmethoden

Für den Effekt der Schultart auf den Einsatz schülerorientierter Unterrichtsmethoden können drei Mediatoren identifiziert werden. Die Mediatoren werden zunächst in separaten Modellen analysiert. Anschließend werden die Mediatoren in einem gemeinsamen Modell betrachtet.

Der Effekt der Schultart auf schülerorientierte Unterrichtsmethoden wird mediert über das Merkmal mathematische Konzepte, wie die Ergebnisse in Abbildung 19 indizieren. Der Effekt der Schultart auf schülerorientierte Unterrichtsmethoden liegt nicht mehr vor, wenn die Mediatorvariable ins Modell aufgenommen wird. Die Pfade β_{RS} , β_{IGS} und β_{MBG} sind also nicht mehr signifikant. Die Schultart hängt vielmehr negativ mit dem Merkmal mathematische Konzepte zusammen ($\beta_{RS_M}=-0,63^*$; $\beta_{IGS_M}=-0,80^*$; $\beta_{MBG_M}=-0,77^*$), welches wiederum negativ mit schülerorientierten Unterrichtsmethoden zusammenhängt ($\beta_M=-0,25^*$). Dies bedeutet, dass in Realschulen, integrierten Gesamtschulen und Schulen mit mehreren Bildungsgängen das Ausmaß, in dem mathematische Konzepte vermittelt werden, geringer ist als an Gymnasien. Dies wurde bereits ausführlich in den deskriptiven Analysen beschrieben. Je geringer das Ausmaß ist, in dem mathematische Konzepte vermittelt werden, desto häufiger werden

schülerorientierte Unterrichtsmethoden eingesetzt. Dadurch ergibt sich ein positiver indirekter Effekt der Schulart auf schülerorientierte Unterrichtsmethoden, welcher über das Merkmal mathematische Konzepte vermittelt wird ($\beta_{RS_IND}=0,16^*$; $\beta_{IGS_IND}=0,20^*$; $\beta_{MBG_IND}=0,12^*$).

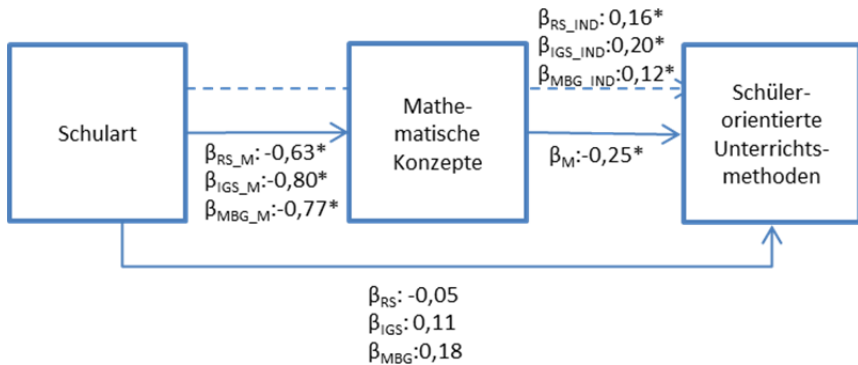


Abbildung 19. Mediation der Schulartunterschiede in schülerorientierten Unterrichtsmethoden über mathematische Konzepte.

Signifikante Effekte sind mit einem Stern (*) gekennzeichnet.

Abkürzungen: MBG=Schulen mit mehreren Bildungsgängen, RS=Realschulen, IGS=integrierte Gesamtschulen, M=Mediator; IND=indirekter Effekt ---->.

Als weiterer Mediator der Schulartunterschiede in den schülerorientierten Unterrichtsmethoden wird das Merkmal reine Mathematik, also die Häufigkeit von rein mathematischen Aufgaben im Unterricht, untersucht (Abbildung 20). Der Effekt der Schulart auf schülerorientierte Unterrichtsmethoden liegt nicht mehr vor, wenn die Mediatorvariable ins Modell aufgenommen wird. Die Pfade β_{RS} , β_{IGS} und β_{MBG_M} sind also nicht mehr signifikant. Die Schulart hängt vielmehr negativ mit dem Merkmal reine Mathematik zusammen ($\beta_{RS_M}=-0,34^*$; $\beta_{IGS_M}=-0,51^*$; $\beta_{MBG_M}=-0,46^*$), welches wiederum negativ mit dem Potential zur kognitiven Aktivierung zusammenhängt ($\beta_M=-0,27^*$). Dies bedeutet, dass in Realschulen, integrierten Gesamtschulen und Schulen mit mehreren Bildungsgängen die Häufigkeit rein mathematischer Aufgaben geringer ist als an Gymnasien. Dies wurde bereits ausführlich in den deskriptiven Analysen beschrieben. Je geringer die Häufigkeit rein mathematischer Aufgaben ist, desto häufiger werden schülerorientierte Unterrichtsmethoden eingesetzt. Dadurch ergibt sich ein positiver indirekter Effekt der Schulart auf schülerorientierte

Unterrichtsmethoden, welcher über das Merkmal reine Mathematik vermittelt wird ($\beta_{RS_IND}=0,09^*$; $\beta_{IGS_IND}=0,14^*$; $\beta_{MBG_IND}=0,13^*$).

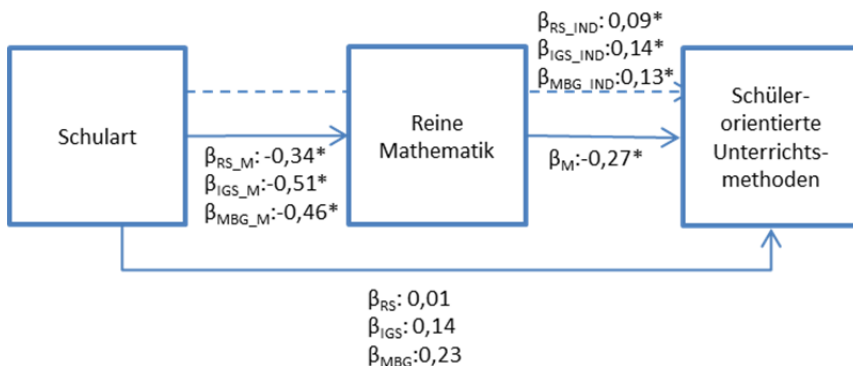


Abbildung 20. Mediation der Schulartunterschiede in schülerorientierten Unterrichtsmethoden über reine Mathematik.

Signifikante Effekte sind mit einem Stern (*) gekennzeichnet.

Abkürzungen: MBG=Schulen mit mehreren Bildungsgängen, RS=Realschulen, IGS=integrierte Gesamtschulen, M=Mediator; IND=indirekter Effekt ---->.

Als drittes Merkmal erweist sich auch die Leistungskomposition als Mediatorvariable für die Schulartunterschiede im schülerorientierten Unterricht (Abbildung 21). Der Effekt der Schulart auf schülerorientierte Unterrichtsmethoden liegt nicht mehr vor, wenn die Mediatorvariable ins Modell aufgenommen wird. Die Pfade β_{RS} , β_{IGS} und β_{MBG} sind also nicht mehr signifikant. Die Schulart hängt vielmehr negativ mit der Leistungskomposition zusammen ($\beta_{RS_M}=-1,12^*$; $\beta_{IGS_M}=-1,32^*$; $\beta_{MBG_M}=-1,44^*$), welche wiederum negativ mit schülerorientierten Unterrichtsmethoden zusammenhängt ($\beta_M=-0,38^*$). Dies bedeutet, dass in Realschulen, integrierten Gesamtschulen und Schulen mit mehreren Bildungsgängen die mittlere Klassenleistung im PISA-Test geringer ist als an Gymnasien. Je geringer die mittlere Leistung ist, desto häufiger werden schülerorientierte Unterrichtsmethoden eingesetzt. Dadurch ergibt sich ein positiver indirekter Effekt der Schulart auf schülerorientierte Unterrichtsmethoden, welcher über die Leistungskomposition vermittelt wird ($\beta_{RS_IND}=0,43^*$; $\beta_{IGS_IND}=0,50^*$; $\beta_{MBG_IND}=0,55^*$).

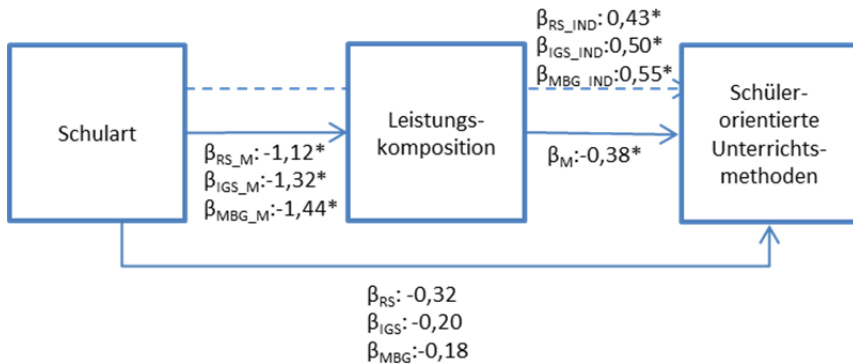


Abbildung 21. Mediation der Schulartunterschiede in schülerorientierten Unterrichtsmethoden über die mittlere Klassenleistung im PISA-Mathematiktest. Signifikante Effekte sind mit einem Stern (*) gekennzeichnet. Abkürzungen: MBG=Schulen mit mehreren Bildungsgängen, RS=Realschulen, IGS=integrierte Gesamtschulen, M=Mediator; IND=indirekter Effekt ---->.

Für die Schulartunterschiede in den schülerorientierten Unterrichtsmethoden können zusammenfassend drei Mediatoren identifiziert werden: mathematische Konzepte, Leistungskomposition und reine Mathematik. In einem gemeinsamen Modell sollen abschließend die drei Mediatoren gemeinsam aufgenommen werden. Abbildung 22 zeigt das zugehörige Analysemodell.

Im Modell mit drei Mediatoren werden lediglich die direkten Effekte der Schulart auf die Mediatoren signifikant. Alle weiteren direkten und indirekten Effekte sind nicht signifikant. Es ist zu vermuten, dass das den teils hohen Interkorrelationen der Merkmale geschuldet ist.

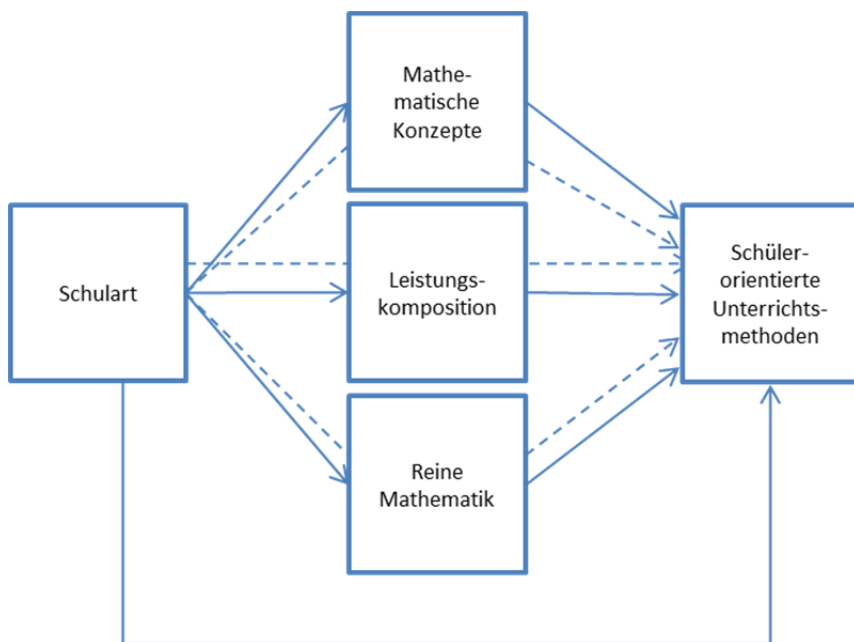


Abbildung 22. Mediationsmodell der Schulartunterschiede in schülerorientierten Unterrichtsmethoden über die Merkmale mathematische Konzepte, Leistungskomposition und reine Mathematik.

Gestrichelte Pfeile kennzeichnen die indirekten Effekte ---->.

7.3.3 Zusammenfassende Interpretation

Ziel der Analysen zu Fragestellung 3 war es, die theoretisch hergeleiteten Mechanismen für Schulartunterschiede in Unterrichtsprozessmerkmalen empirisch zu prüfen. Die Analysen waren dabei beschränkt auf vier Unterrichtsprozessmerkmale (abhängige Variablen): Schülerorientierte Unterrichtsmethoden, formatives Assessment, Potential zur kognitiven Aktivierung durch Begründen und Potential zur kognitiven Aktivierung durch Variation der Aufgaben. Andere Unterrichtsprozessmerkmale wurden aufgrund mangelnder Vergleichbarkeit oder ausbleibender Schulartunterschiede nicht in die Analysen einbezogen.

Für die Schulartunterschiede von zwei Unterrichtsprozessmerkmalen können vermittelnde Mechanismen identifiziert werden: Für schülerorientierte Unter-

richtsmethoden und das Potential zur kognitiven Aktivierung durch Variation der Aufgaben werden jeweils drei mögliche Arten von Mediatoren unterschieden a) Überzeugungen der Lehrenden, b) Unterrichtsinhalte und c) die Zusammensetzung der Schülergruppe. Auf Seiten der Lehrenden mediierten die analysierten Überzeugungen der Lehrenden nicht die Schulartunterschiede in den Unterrichtsprozessmerkmalen. Für die untersuchten Merkmale lässt sich damit vorsichtig folgern, dass die Überzeugungen der Lehrenden nicht bedeutend sind für Schulartunterschiede in den eingesetzten Unterrichtsmethoden oder dem Potential zur kognitiven Aktivierung.

Die Unterrichtsinhalte scheinen hingegen von größerer Bedeutung für Unterrichtsprozesse zu sein. Beide Merkmale, welche Unterrichtsinhalte beschreiben, können Unterschiede in den Unterrichtsprozessen erklären. Ein stärkerer Fokus auf die Vermittlung von Fachkonzepten in Mathematik geht einher mit einem geringeren Einsatz schülerorientierter Unterrichtsmethoden. Es ist zu vermuten, dass Fachkonzepte eher in Form von Frontalunterricht vermittelt werden. Auch ein häufigerer Einsatz rein mathematischer Aufgaben scheint den Einsatz schülerorientierter Unterrichtsmethoden zu minimieren. Da an Realschulen, integrierten Gesamtschulen und Schulen mit mehreren Bildungsgängen diese Inhalte seltener Teil des Mathematikunterrichts sind als an Gymnasien, werden schülerorientierte Unterrichtsmethoden somit häufiger im Unterricht eingesetzt. Gleichzeitig ist das Potential zur kognitiven Aktivierung durch die Variation der Aufgaben höher, wenn der Fokus auf die Vermittlung von mathematischen Konzepten stärker ist. Die Lehrenden prüfen vermutlich anhand der variierenden Aufgaben, ob die mathematischen Konzepte verstanden wurden. Mediiert über die Vermittlung von mathematischen Konzepten ist somit das Potential zur kognitiven Aktivierung durch die Variation der Aufgaben an Realschulen, integrierten Gesamtschulen und Schulen mit mehreren Bildungsgängen geringer als an Gymnasien.

Die Schülergruppe scheint ebenfalls Schulartunterschiede im Einsatz schülerorientierter Unterrichtsmethoden zu mediierten. Lehrende setzen bei einer geringeren mittleren Klassenleistung häufiger schülerorientierte Unterrichtsmethoden ein. Entsprechend der unterschiedlichen mittleren Leistungszusammensetzung zwischen den Schularten werden schülerorientierte Unterrichtsmethoden in Realschulen, integrierten Gesamtschulen und Schulen mit mehreren Bildungsgängen häufiger eingesetzt als an Gymnasien. Hier stellt sich die Frage nach der Kausalität. Auf der einen Seite könnten die Schülerinnen und Schüler

in Klassen mit niedrigerer mittlerer Mathematikleistung tatsächlich besser lernen, wenn die Unterrichtsmethoden stärker schülerorientiert sind. Auf der anderen Seite könnten Lehrende aufgrund von Überzeugungen und Vorurteilen im Sinne des Pygmalioneffektes gegenüber der Klasse handeln (siehe im Theorie-Teil Abschnitt 3.4.5), ohne dass eine stärkere Schülerorientierung in den Methoden angemessen wäre. Auf diesen Überlegungen baut Fragestellung 4 auf, welche im nächsten Abschnitt empirisch bearbeitet wird.

7.4 Schultartunterschiede in den Zusammenhängen zwischen Unterrichtsprozessen und Leistungsentwicklung: Moderationseffekte der Schultart

Die bisherigen Analysen kamen zu dem Ergebnis, dass die Klassen der verschiedenen Schularten bedeutende Unterschiede in den Unterrichtsprozessen bezogen auf schülerorientierten und kognitiv aktivierenden Unterricht sowie im Einsatz formativen Assessments aufweisen. Ferner werden diese Unterschiede unter anderem über Unterrichtsinhalte und die kognitiven Eigenschaften der Schülergruppe vermittelt. Die gefundenen Unterschiede in den Unterrichtsprozessen und den vermittelnden Mechanismen sind konform mit den intendierten Schwerpunkten der Schularten, wie sie in Kapitel 2 herausgearbeitet wurden.

Wie keine andere Schultart betont das Gymnasium den fachlichen Zugang sowie eine starke Reflexionskultur, was sich in den deskriptiven Analysen im vergleichsweise hohen Potential zur kognitiven Aktivierung im Unterricht widerspiegelt. In den Hintergrund treten an dieser Schultart hingegen das Erfahrungslernen sowie erzieherische Ansprüche; Ziele, die an den anderen Schularten stärker im Fokus stehen. Besonders Gesamtschulen betonen die Orientierung am Lernenden bzw. an der Lernenden und die Förderung individueller Interessen und Fähigkeiten. Dies spiegelt sich in den vergleichsweise hohen Ausprägungen bei den Merkmalen schülerorientierte Unterrichtsmethoden und formatives Assessment bei integrierten Gesamtschulen und Schulen mit mehreren Bildungsgängen wider.

Abschließend wird nun der Frage nachgegangen, ob die Zusammenhänge zwischen den Unterrichtsprozessen und der Leistungsentwicklung in allen Schularten

ten gleich sind, oder ob die Effekte der Unterrichtsprozessmerkmale auf die Leistungsentwicklung von der Schulart abhängig sind. Zur Beantwortung der Frage wird analysiert, ob die Zusammenhänge zwischen jeweils einem Unterrichtsprozessmerkmal und der Leistungsentwicklung durch die Schulart beeinflusst bzw. moderiert werden, d.h., es wird untersucht, ob die Schulart mit dem Unterrichtsprozessmerkmal interagiert. Gemessen wird der moderierende Effekt der Schulart, indem zur Vorhersage der Leistung zum zweiten Messzeitpunkt neben dem Unterrichtsprozessmerkmal und der Schulart zusätzlich der Interaktionsterm als Produkt dieser beiden Merkmale aufgenommen wird. Finden sich signifikante Effekte des Interaktionsterms (Interaktionseffekte), so wäre nicht davon auszugehen, dass die Zusammenhänge zwischen einem Unterrichtsprozessmerkmal und der Leistungsentwicklung generalisiert werden können, wie häufig in der Schuleffektivitätsforschung angenommen wird. Vielmehr wäre anzunehmen, dass die Unterrichtsprozesse abhängig von der Schulart die Entwicklung der Lernenden unterschiedlich beeinflussen.

Analysiert werden vier Unterrichtsprozessmerkmale – schülerorientierte Unterrichtsmethoden, formatives Assessment, kognitive Aktivierung durch Variation und kognitive Aktivierung durch Begründen –, für welche die deskriptiven Analysen bedeutende Schulartunterschiede indizieren (Abschnitt 7.2.7), und deren Zusammenhänge mit der Leistungsentwicklung. Die Leistungsentwicklung wird zum einen gemessen über den PISA-Test und zum anderen über den BiSta-Test. Somit werden für jedes Unterrichtsprozessmerkmal jeweils zwei Modelle berechnet und die Interaktionseffekte getestet. Wie die Analysen in Abschnitt 7.1.2 gezeigt haben, sind die Antworten der Lernenden und der Lehrenden durch Antworttendenzen verzerrt, weshalb die Modelle unter Kontrolle dieser Antworttendenzen berechnet werden. Die Leistungsentwicklung wird modelliert, indem auf Individualebene die Leistung im zweiten Erhebungszeitpunkt kontrolliert wird durch die Leistung zum ersten Erhebungszeitpunkt (siehe im Methodenteil Abschnitt 6.9). Im Folgenden sind die Ergebnisse zunächst für den BiSta-Test dargestellt, anschließend für den PISA-Test. Abschnitt 7.4.3 fasst die zentralen Befunde zusammen.

7.4.1 Zusammenhänge von Unterrichtsprozessmerkmalen und Leistungsentwicklung gemessen mit dem BiSta-Test

Tabelle 18 gibt die Haupteffekte der Unterrichtsprozessmerkmale und der Schulart sowie die Interaktionseffekte auf die Leistung zum zweiten Erhebungszeitpunkt gemessen mit dem BiSta-Test wieder. Signifikante Haupteffekte der Schulart liegen für alle Unterrichtsprozessmerkmale vor, was bedeutet, dass Klassen der Realschulen, integrierte Gesamtschulen und Schulen mit mehreren Bildungsgängen im Durchschnitt signifikant geringere Leistungen im BiSta-Test des zweiten Erhebungszeitraums aufzeigen als Klassen an Gymnasien (bei einem Signifikanzniveau von $p=0,05$), wenn für die Leistung zum ersten Erhebungszeitraum auf Individualebene kontrolliert wird.

Bei den Unterrichtsprozessmerkmalen kann nur für das Merkmal „schülerorientierte Unterrichtsmethoden“ ein signifikanter Zusammenhang mit der Leistung zum zweiten Erhebungszeitraum gefunden werden. Für Klassen an Gymnasien ist der Zusammenhang zwischen schülerorientierten Unterrichtsmethoden und der Leistung zum zweiten Erhebungszeitpunkt negativ ($\beta_{UP}=-0,10^*$). Unter Kontrolle der individuellen Mathematikleistungen zum ersten Erhebungszeitpunkt ist daher davon auszugehen, dass ein häufigerer Einsatz schülerorientierter Unterrichtsmethoden in Klassen an Gymnasien negativ mit der Leistung zum zweiten Erhebungszeitpunkt zusammenhängt. Dieser Zusammenhang gilt nicht uneingeschränkt für alle Schularten, was die signifikanten Interaktionsterme für Realschulen und Schulen mit mehreren Bildungsgängen indizieren. Für Klassen an Realschulen fällt der negative Zusammenhang deutlich niedriger aus ($\beta_{UP} + \beta_{RS*UP} = -0,10 + 0,07 = -0,03$). Für Klassen an Schulen mit mehreren Bildungsgängen ist der Zusammenhang leicht positiv ($\beta_{UP} + \beta_{MBG*UP} = -0,10 + 0,17 = 0,07$). An Schulen mit mehreren Bildungsgängen profitieren die Lernenden scheinbar von schülerorientierten Unterrichtsmethoden. Schülerorientierte Unterrichtsmethoden haben demnach abhängig von der Schulart unterschiedliche Effekte auf die Leistungen bzw. die Leistungsentwicklung der Lernenden im BiSta-Test.

Tabelle 18. Moderierende Effekte der Schulart auf den Zusammenhang von Unterrichtsprozessmerkmalen und der Leistungsentwicklung gemessen mit dem BiSta-Test.

Leistungsentwicklung in BiSta		Schüler-orientierte Unterrichtsmethoden	Formatives Assessment	Kognitive Aktivierung durch Begründen	Kognitive Aktivierung durch Variation
Schüler-ebene	BiSta T1	β 0,75 *	β 0,75 *	β 0,75 *	β 0,76 *
	R ²	0,61 *	0,61 *	0,61 *	0,62 *
Klassen-ebene	Intercept	0,06 *	0,07 *	0,07 *	0,07 *
	UP	-0,10 *	-0,05	0,03	-0,01
	RS	-0,14 *	-0,15 *	-0,10 *	-0,11 *
	IGS	-0,15 *	-0,15 *	-0,15 *	-0,16 *
	MBG	-0,16 *	-0,12 *	-0,12 *	-0,13 *
	RS*UP	0,07 *	0,06	0,02	0,06
	IGS*UP	0,09	0,05	-0,02	-0,05
	MBG*UP	0,17 *	0,04	-0,01	-0,05
	ARS	0,05 *	0,04	0,01	-0,01
	MPR	0,01	0,01	0,01	0,01
	R ²	0,86 *	0,76	0,78	0,85

Die Kontrollvariablen und die zugehörigen Ergebnisse sind grau markiert.

Signifikante Effekte ($p < 0,05$) sind mit einem Stern (*) gekennzeichnet.

Abkürzungen: UP=Unterrichtsprozessmerkmal; RS=Realschule;

IGS=integrierte Gesamtschule; MBG=Schulen mit mehreren Bildungsgängen.

7.4.2 Zusammenhänge von Unterrichtsprozessmerkmalen und Leistungsentwicklung gemessen mit dem PISA-Test

Tabelle 19 gibt die Haupteffekte der Unterrichtsprozessmerkmale und der Schulart sowie die Interaktionseffekte auf die Leistung zum zweiten Erhebungszeitpunkt gemessen mit dem PISA-Test wieder. Auch hier liegen signifikante Haupteffekte der Schulart für alle Unterrichtsprozessmerkmale vor, was bedeutet, dass Klassen der Realschulen, integrierte Gesamtschulen und Schulen mit mehreren Bildungsgängen im Durchschnitt signifikant geringere Leistungen im PISA-Test des zweiten Erhebungszeitraums aufzeigen als Klassen an Gymnasien (bei einem Signifikanzniveau von $p = 0,05$), wenn für die Leistung zum ersten Erhebungszeitraum auf Individualebene kontrolliert wird.

Bei den Unterrichtsprozessmerkmalen können für zwei Merkmale – schülerorientierte Unterrichtsmethoden und formatives Assessment – signifikante Zusammenhänge mit der Leistung zum zweiten Erhebungszeitraum gefunden werden. Für Klassen an Gymnasien sind die Zusammenhänge im Mittel negativ ($\beta_{UP} = -0,04^*$ bzw. $\beta_{UP} = 0,05^*$). Unter Kontrolle der individuellen Mathematikleistungen zum ersten Erhebungszeitpunkt ist daher davon auszugehen, dass ein häufigerer Einsatz schülerorientierter Unterrichtsmethoden und formativen Assessments in Klassen an Gymnasien zu einer geringeren Leistung im PISA-Test des zweiten Erhebungszeitpunktes führen. Für schülerorientierte Unterrichtsmethoden gilt dieser Zusammenhang wiederum nicht uneingeschränkt für alle Schularten, was der signifikante Interaktionsterm für integrierte Gesamtschulen indiziert. Für Klassen dieser Schulart fällt der Zusammenhang im Mittel positiv aus ($\beta_{UP} + \beta_{IGS*UP} = -0,04 + 0,08 = 0,04$). Für Klassen an Realschulen und Schulen mit mehreren Bildungsgängen weichen die Zusammenhänge nicht von Ergebnissen für das Gymnasium ab. Formatives Assessment scheint im Mittel für Klassen aller Schularten negativ mit der Leistungsentwicklung zusammenzuhängen.

Tabelle 19. Moderierende Effekte der Schulart auf den Zusammenhang von Unterrichtsprozessmerkmalen und der Leistungsentwicklung gemessen mit dem PISA-Test.

Leistungsentwicklung in PISA		Schüler-orientierte Unterrichtsmethoden	Formatives Assessment	Kognitive Aktivierung durch Begründen	Kognitive Aktivierung durch Variation
		β	β	β	β
Schüler-ebene	PISA T1	0,86 *	0,86 *	0,84 *	0,84 *
	R ²	0,82 *	0,82 *	0,80 *	0,80 *
Klassen-ebene	Intercept	0,07 *	0,06 *	0,10 *	0,10 *
	UP	-0,04 *	-0,05 *	0,00	0,00
	RS	-0,11 *	-0,10 *	-0,18 *	-0,17 *
	IGS	-0,19 *	-0,17 *	-0,17 *	-0,17 *
	MBG	-0,22 *	-0,23 *	-0,21 *	-0,19 *
	RS*UP	-0,01	0,01	-0,02	-0,03
	IGS*UP	0,08 *	0,06	0,03	-0,03
	MBG*UP	0,03	0,06	0,04	0,17
	ARS	0,04 *	0,04 *	0,01	0,01
	MPR	-0,03 *	-0,03 *	0,00	0,00
	R ²	0,46 *	0,45 *	0,51 *	0,60 *

Die Kontrollvariablen und die zugehörigen Ergebnisse sind grau markiert.

Signifikante Effekte ($p < 0,05$) sind mit einem Stern (*) gekennzeichnet.

Abkürzungen: UP=Unterrichtsprozessmerkmal; RS=Realschule;

IGS=integrierte Gesamtschule; MBG=Schulen mit mehreren Bildungsgängen.

7.4.3 Zusammenfassende Interpretation

Die Unterschiede in den Unterrichtsprozessen zwischen den Schularten wurden in diesem Ergebnisteil hinsichtlich ihrer Bedeutung für die Leistungsentwicklung der Lernenden im BiSta-Test und im PISA-Test untersucht. Insbesondere wurde der Frage nachgegangen, ob die Zusammenhänge zwischen den Unterrichtsprozessmerkmalen und der Leistungsentwicklung zwischen den Schularten Unterschiede aufweisen. Es zeigt sich, dass die Unterrichtsprozessmerkmale, welche das Potential zur kognitiven Aktivierung durch Variation und das Potential zur kognitiven Aktivierung durch Begründen messen, nicht mit der Leistungsentwicklung der Klassen zusammenhängen. Die Merkmale spiegeln

die Absichten der Lehrenden wider und sind kein Garant, dass diese im Unterricht auch umgesetzt wurden (Kunter und Trautwein 2013).

Die Merkmale, welche schülerorientierte Unterrichtsmethoden sowie formatives Assessment messen, scheinen hingegen mit der Leistungsentwicklung zusammenzuhängen. Wobei sich die Zusammenhänge zwischen den Schularten unterscheiden. Für Klassen an Gymnasien sind die Zusammenhänge generell negativ. Für Klassen an Realschulen scheint sich der negative Zusammenhang zwischen formativem Assessment und der Leistung zum zweiten Erhebungszeitpunkt fast aufzuheben. Konträre Ergebnisse zeigen sich für Klassen an Schulen mit mehreren Bildungsgängen, für welche der Einsatz schülerorientierter Unterrichtsmethoden positiv mit der Leistungsentwicklung im BiSta-Test zusammenhängt. Klassen an integrierten Gesamtschulen zeigen zwar nicht im BiSta-Test, jedoch im PISA-Test bessere Leistungen, wenn häufiger schülerorientierte Unterrichtsmethoden eingesetzt werden.

Unterrichtsprozesse, hier speziell schülerorientierte Unterrichtsmethoden, können demnach je nach Schulart unterschiedlich geeignet sein, um Lernende in ihrer Leistungsentwicklung zu fördern. Auffällig ist jedoch, dass die Ergebnisse zwischen den Leistungstests abweichen. Diese Befunde werden in der Diskussion wieder aufgegriffen. Hier sei nur kurz darauf verwiesen, dass die eingesetzten Leistungstests unterschiedliche Schwerpunkte mathematischen Wissens und Könnens erheben. Die eingesetzten Unterrichtsmethoden könnten demnach in der Stichprobe je nach Schulart unterschiedliche Bereiche mathematischen Wissens und Könnens adressiert bzw. gefördert haben.

8 Diskussion

Der vorliegende Beitrag zur Schulforschung verbindet die Ebene der Schulstruktur mit der Unterrichtsebene. Konkret wird hinterfragt, welche Bedeutung das gegliederte Bildungssystem der Sekundarstufe I für die Prozesse im Unterricht hat. Untersucht wird diese Frage, indem Unterrichtsprozesse vergleichend zwischen den Schularten betrachtet und Unterschiede herausgearbeitet werden.

Aus der historischen Entwicklung heraus unterscheiden sich die Schularten in ihren Leitbildern und Anschlussperspektiven sowie in den Bildungszielen und -inhalten. Zudem unterscheidet sich die Gruppe der Lernenden; zum einen durch die teils intendierte Leistungsverteilung, zum anderen durch Segregationsprozesse (vgl. Maaz et al. 2010). Darüber hinaus finden sich bis heute Unterschiede in der Lehrerbildung an den Hochschulen und im anschließenden Studienseminar (Abs 2006; Blömeke 2019). Diese systembedingten Unterschiede zwischen den Schularten wirken sich, so die Hypothese dieser Arbeit, auf den Unterricht aus. In Anlehnung an das Modell des Didaktischen Dreiecks wird in dieser Arbeit der Unterricht als ein relationales und interdependentes Geschehen zwischen eben diesen Komponenten gesehen: der Lehrperson, der Schülergruppe und den Unterrichtsinhalten. Hieraus begründet sich die Annahme dieser Arbeit, dass durch die systembedingten Unterschiede zwischen den Schularten, welche sich unmittelbar auf die Lehrenden, die Schülergruppe und die Unterrichtsinhalte auswirken, Unterschiede im Unterricht bzw. in den damit verbunden Unterrichtsprozessen zwischen den Schularten vorliegen.

In Kapitel 3 wurde tiefergehend begründet, wie und warum sich die Überzeugungen der Lehrenden, die Unterrichtsinhalte sowie die Gruppe der Lernenden zwischen den Schularten unterscheiden. Zudem wurden Argumente dafür aufgeführt, weshalb diese drei Merkmalsbereiche für die unterrichtlichen Prozesse von Bedeutung sein könnten. Der abschließende Forschungsüberblick zu Schulartunterschieden verdeutlichte schließlich auch empirisch die systematischen Unterschiede hinsichtlich der Überzeugungen der Lehrenden, der Zusammensetzung der Schülergruppe und der Unterrichtsinhalte sowie systematische Unterschiede in den Unterrichtsprozessen. Gleichzeitig konnten Forschungslücken identifiziert werden. Zum einen werden bei den deskriptiven Vergleichen anderer Publikationen häufig keine Signifikanztests berichtet, welche die Varianz innerhalb und zwischen den Schularten berücksichtigen. Zum anderen werden nur selten die Schularten Realschule, Schulen mit mehre-

ren Bildungsgängen und integrierte Gesamtschulen in der Forschung separat analysiert oder beschrieben. Besonders für diese Schularten fehlen daher vergleichende Informationen. Schließlich wurden in den Analysen, welche im Forschungsüberblick zusammengefasst wurden, wichtige Vorannahmen bezüglich der Vergleichbarkeit der Daten anscheinend nicht geprüft. Durch systematische Unterschiede zwischen den Lernenden sowie zwischen den Lehrenden, könnte die Interpretation einzelner Konstrukte zwischen den Befragten der verschiedenen Schularten variieren. Zudem könnten Fragebogendaten durch Antworttendenzen systematisch verzerrt sein.

Basierend auf den oben beschriebenen Grundannahmen bezüglich systematischer Schulartunterschiede und den identifizierten Forschungsdesideraten wurden in Kapitel 5 die Fragestellungen der Arbeit expliziert und die hierfür verwendeten Daten und Methoden in Kapitel 6 beschrieben. Im empirischen Teil der Arbeit (Kapitel 7) wurde grundlegend die Vergleichbarkeit von Fragebogendaten zwischen den Schularten geprüft (Abschnitt 7.1). Darauf aufbauend wurden die Schularten anhand der für den Vergleich geeigneten Daten und unter Kontrolle von Antworttendenzen deskriptiv gegenübergestellt und die Übereinstimmung mit Befunden aus anderen Schulartvergleichen expliziert (Abschnitt 7.2). Anschließend wurde geprüft, inwieweit die theoretisch hergeleiteten Vermittlungsmechanismen für schulartspezifische Unterrichtsprozesse – nämlich Mediation durch Überzeugungen der Lehrenden, Zusammensetzung der Schülergruppe und Unterrichtsinhalte – empirisch nachgewiesen werden können (Abschnitt 7.3). Abschließend wurde der Frage nachgegangen, inwieweit sich nicht nur die Unterrichtsprozesse, sondern auch deren Zusammenhänge mit der Leistungsentwicklung der Lernenden zwischen den Schularten unterscheiden (Abschnitt 7.4).

Die folgende Diskussion fasst die empirischen Ergebnisse und theoretischen Erkenntnisse auf einer übergeordneten Ebene zusammen. Zudem beleuchtet die Diskussion Implikationen für weitere Forschungen sowie Überlegungen zu den Schularten und zur Schulstruktur in der Sekundarstufe I. Ebenso werden die Grenzen der vorliegenden Arbeit sowie mögliche Hinweise für die Praxis reflektiert. Konkret werden die Kernergebnisse der Arbeit zu den deskriptiven Schulartunterschieden und zu den möglichen Zusammenhängen mit der Leistungsentwicklung der Lernenden dargestellt und diskutiert (Abschnitt 8.1). In Abschnitt 8.1 werden die Erkenntnisse zur Vergleichbarkeit von Schularten verallgemeinert und Möglichkeiten besprochen, die Vergleichbarkeit von Fra-

gebogendaten in der Schul- und Unterrichtsforschung abzusichern. Anschließend wird auf mögliche Kritikpunkte und Grenzen des empirischen Ansatzes sowie auf die Grenzen der verwendeten Methoden und Daten eingegangen (Abschnitt 8.3). Die beiden letzten Abschnitte besprechen die empirischen Erkenntnisse im Kontext der Schulstrukturdebatte: Abschnitt 8.4 stellt die Bedeutung der Einzelschule der Schulart gegenüber. Abschnitt 8.5 diskutiert abschließend die Schularten und die Schulstruktur im Kontext gesellschaftlicher, wirtschaftlicher und individueller Anforderungen.

8.1 Bedeutung der Schulartunterschiede

In diesem Abschnitt werden die zentralen Befunde der deskriptiven Beschreibung der Schularten sowie schulartspezifische Zusammenhänge von Unterrichtsprozessen und Mathematikleistung gemeinsam diskutiert. Abbildung 23 fasst die Befunde für Merkmale, in denen sich die Schularten bedeutend unterscheiden, verdichtet zusammen.²² Dabei werden alle Merkmale in drei Stufen klassifiziert. Ein großer Balken (■) bedeutet, dass das entsprechende Merkmal im Vergleich zu den anderen Schularten hoch ausgeprägt ist, ein kleiner Balken (▀) gibt an, dass die Merkmalsausprägung im Mittelfeld liegt. Ein Strich (—) bedeutet, dass das entsprechende Merkmal in der Schulart vergleichsweise gering ausgeprägt ist. Der Zusammenhang der Unterrichtsprozessmerkmale mit der Leistung, unter Kontrolle der Eingangsleistung, wird zusammenfassend als positiv (+) oder negativ (–) dargestellt, falls er signifikant ist ($p < 0,05$). Signifikante Unterschiede zwischen den Schularten – Schulen mit mehreren Bildungsgängen, integrierte Gesamtschule und Realschule – gegenüber dem Gymnasium sind zusätzlich grau unterlegt (■). Leere Zellen bedeuten, dass kein signifikanter Zusammenhang zwischen dem Merkmal und der Leistung nachweisbar war.

22 Auf die Darstellung der deskriptiven Beschreibung der Schülergruppe hinsichtlich der Mathematikleistung und des sozioökonomischen Status wurde an dieser Stelle verzichtet. Die mehrfach replizierten Unterschiede zwischen den Schularten hinsichtlich dieser Merkmale, die sich auch in den hier verwendeten Daten zeigen, werden als bekannt vorausgesetzt und können der Diskussion keine Impulse geben.

	Deskriptive Beschreibung der Schularten							Zusammenhang mit Leistung							
	Unterrichtsprozesse				Inhalte		Unterrichtsziel	BiSta	PISA	BiSta	PISA	BiSta	PISA	BiSta	PISA
	Schülerorientierte Unterrichtsmethoden	Formatives Assessment	Kognitive Aktivierung durch Begründen	Kognitive Aktivierung durch Variation	Reine Mathematik	Mathematische Konzepte	Modellierungsfähigkeit	Schülerorientierte Unterrichtsmethoden	Formatives Assessment	Kognitive Aktivierung durch Begründen	Kognitive Aktivierung durch Variation				
MBG								+	-		-				
IGS								-	+		-				
RS								-	-		-				
GY								-	-		-				

Abbildung 23. Zusammenfassung der Schulartunterschiede.

Unterscheidung der Merkmalsausprägungen: Großer Balken = Merkmal im Vergleich zu den anderen Schularten hoch ausgeprägt. Kleiner Balken = Merkmal im Vergleich zu den anderen Schularten im Mittelfeld. Strich = Merkmal im Vergleich zu den anderen Schularten gering ausgeprägt.

Abkürzungen: MBG=Schule mit mehreren Bildungsgängen, IGS=Integrierte Gesamtschule; RS=Realschule, GY=Gymnasium.

Die Gegenüberstellung der Schularten zeigt deutlich den Schwerpunkt der integrierten Gesamtschulen und Schulen mit mehreren Bildungsgängen in einer hohen Schülerorientierung. Schülerorientierte Unterrichtsmethoden und formatives Assessment finden in diesen Schularten vergleichsweise häufig statt. Überraschend ist, dass formatives Assessment an integrierten Gesamtschulen seltener umgesetzt wird als an Schulen mit mehreren Bildungsgängen. Dabei wird vor allem diese Schulart mit einer starken Orientierung am Individuum verbunden (siehe Kapitel 2) und die integrierte Gesamtschule als „Schule der optimalen individuellen Persönlichkeits- und Leistungsentwicklung“ auch heute noch durch die Gemeinnützige Gesellschaft für Gesamtschulen beworben (Gemeinnützige Gesellschaft Gesamtschule 2015). Es wäre daher zu vermuten, dass formatives Assessment, neben schülerorientierten Unterrichtsmethoden, an integrierten Gesamtschulen vergleichsweise häufig realisiert wird. Denn besonders durch formatives Assessment können Entwicklungs- bzw. Lernspielräume identifiziert und damit der Lernende in seinen Entwicklungs- und Lernprozessen unterstützt werden (siehe hierfür die Schlüsselmerkmale formativen Assessments nach Wiliam und Thompson (2008), beschrieben in Abschnitt 3.1.2 dieser Arbeit).

Ungeklärt bleibt an dieser Stelle, ob Idealbild und Realität nicht ganz übereinstimmen oder ob integrierte Gesamtschulen eine vergleichsweise hohe Orien-

tierung an der individuellen Persönlichkeits- und Leistungsentwicklung auf andere als die in dieser Arbeit untersuchten Prozesse erreicht. Beispielsweise ist die Operationalisierung formativen Assessments, wie es in PISA erhoben wurde, stark auf Feedback im Mathematikunterricht bezogen. Somit werden die übergeordneten Ziele der integrierten Gesamtschule, bspw. individuelle Schwerpunktlegung durch die Wahl von Fächern und Kursen, nicht abgefragt. Zudem sei angemerkt, dass es sich hier um ein relatives Abbild handelt. Werden die Rohwerte einbezogen, wird deutlich, dass formatives Assessment durchaus Teil des Unterrichtsalltags an integrierten Gesamtschulen ist. So geben zwischen 30 % und 40 % der Lernenden an integrierten Gesamtschulen an, dass verschiedene Teilaspekte des formativen Assessments (wie Feedback, Erwartungen und Tipps zu individuellen Lernstrategien) in den meisten bzw. in allen Unterrichtsstunden realisiert werden.²³

An Realschulen und besonders an Gymnasien ist der Unterricht weniger schülerorientiert. Sind an Realschulen formatives Assessment, weniger jedoch schülerorientierte Unterrichtsmethoden Teil des Unterrichts, findet an Gymnasien relativ gesehen und auch auf die Rohwerte bezogen schülerorientierter Unterricht nur selten statt. Stattdessen heben sich die Gymnasien durch ein höheres Potential zur kognitiven Aktivierung von den anderen Schularten ab. Damit spiegelt sich empirisch der Anspruch des Gymnasiums als fachliche, kognitive Bildungsanstalt mit starker Reflexionskultur wider. Nur Schulen mit mehreren Bildungsgängen haben ein ähnlich hohes Niveau im Potential zur kognitiven Aktivierung durch die Variation von Aufgaben. Das, zumindest gesetzte, hohe kognitive Anspruchsniveau der Gymnasien spiegelt sich auch in den unterrichteten Inhalten und den Zielen des Mathematikunterrichts wider: Mathematische Konzepte und rein mathematische Aufgaben sind deutlich häufiger Teil des Unterrichts als in den anderen Schularten. Auch das Unterrichtsziel, Mathematisierungsfähigkeiten zu vermitteln, wird hier von den Lehrenden vergleichsweise hoch gewichtet.

Realschulen sind verglichen mit den anderen Schularten bei den herangezogenen Merkmalen eher ausgewogen. Weder in den kognitiv bezogenen Unterrichtsprozessen, Inhalten und Zielen noch in der Schülerorientierung scheint ein Schwerpunkt zu liegen. In der theoretischen Beschreibung (siehe Kapitel 2)

23 Schulen mit mehreren Bildungsgängen: ca. 40–50 %, Realschulen ca. 25–35 %, Gymnasien ca. 10–20 %.

wurden Realschulen als Schulen mit Bezug zum praktischen Leben mit einem Schwerpunkt in der Vorbereitung auf anspruchsvolle Ausbildungsberufe beschrieben. Eventuell schlägt sich dieser Fokus eher im Fächerkanon nieder als konkret in den Zielen, Inhalten oder den Prozessen des Mathematikunterrichts. Anders betrachtet ist es jedoch auch ein Alleinstellungsmerkmal, sowohl in Zielen und Inhalten des Mathematikunterrichts einen mittleren kognitiven Anspruch zu setzen, als auch schülerorientiert zu arbeiten, wenn auch das Potential zur kognitiven Aktivierung in den hier gemessenen Merkmalen gering ist.

Eine starke kognitive Förderung wird in der Schuleffektivitätsforschung als positiv für die Leistungsentwicklung betont. Eine hohe Schülerorientierung wird bezugnehmend auf Deci und Ryan (1996) als förderlich für die intrinsische Motivation angenommen, welche wiederum positiv mit der Leistung assoziiert ist. Deshalb lässt gerade das ausgewogene Angebot an Realschulen eine optimale Entwicklung der Lernenden erwarten. Empirisch zeigt sich jedoch ein anderes Bild. In den meisten Untersuchungen weisen Lernende an Gymnasien die größte Leistungsentwicklung auf (einen Überblick geben Becker et al. 2012). Auch die Daten dieser Arbeit können, gemessen mit dem PISA-Test, diese Ergebnisse replizieren. Gemessen mit dem BiSta-Test ist hingegen die Leistungsentwicklung von Lernenden in integrierten Gesamtschulen und Schulen mit mehreren Bildungsgängen am stärksten. Die konträren Ergebnisse zwischen den Leistungstests sowie die unterschiedlichen Schwerpunkte der Schularten waren unter anderem Gründe, in der hier vorliegenden Arbeit die Zusammenhänge zwischen den Unterrichtsprozessmerkmalen und der Leistungsentwicklung tiefergehend bzw. vergleichend zwischen den Schularten zu betrachten. Die Ergebnisse sind zusammenfassend in Abbildung 23 auf der rechten Seite dargestellt.

Für Merkmale, welche das Potential zur kognitiven Aktivierung messen, können innerhalb der Schularten keine Zusammenhänge mit der Leistung in der 10. Klasse – unter Kontrolle der Mathematikleistung in der 9. Klasse – gefunden werden. Damit können die Ergebnisse dieser Arbeit die theoretisch angenommenen und teils empirisch gefundenen positiven Zusammenhänge zwischen kognitiver Aktivierung und der Leistungsentwicklung nicht bestätigen. Dies könnte auf die verwendeten Merkmalen zurückgeführt werden (siehe hierzu die Diskussion der Limitationen dieser Arbeit in Abschnitt 8.3). Interessanter an dieser Stelle ist jedoch der Befund, dass in den Analysen, welche Zusammenhänge zwischen kognitiver Aktivierung und Leistung finden, vielfach die

Schularten nicht berücksichtigt werden (einen Überblick über den Forschungsstand gibt Lipowsky 2015). Auch mit dem hier vorliegenden Datensatz finden sich teilweise positive Zusammenhänge zwischen der kognitiven Aktivierung und der Leistung in Mathematik (unter Berücksichtigung der Mathematikleistung in der 9. Klasse), wenn die Daten über die Schularten hinweg analysiert werden. Beispielsweise hängt das Merkmal „kognitive Aktivierung durch Begründen“ positiv mit der Leistung im BiSta-Test zusammen. Es liegt daher die Vermutung nahe, dass der positive Zusammenhang ohne Berücksichtigung der Schulart nur besteht, da Lernende an Gymnasium im Mittel höhere Leistungen haben als Lernende anderer Schularten und zugleich die kognitive Aktivierung im Mittel höher ist. Schulart und Leistung sind demnach hoch konfundiert.

Die Bedeutung der Schulart für die Leistungsentwicklung zeigen auch die Ergebnisse für das Merkmal „schülerorientierte Unterrichtsmethoden“, welches in der Stärke sowie in der Richtung des Zusammenhangs mit Leistung teilweise zwischen den Schularten variiert. Schülerorientierte Unterrichtsmethoden scheinen demnach je nach Schulart unterschiedlich gut geeignet zu sein, um Lernende in ihrer Leistungsentwicklung zu fördern. Unterschiede in den Zusammenhängen zeigen sich jedoch nicht nur zwischen den Schularten, sondern auch zwischen den eingesetzten Leistungstests. In Schulen mit mehreren Bildungsgängen zeigen schülerorientierte Unterrichtsmethoden einen positiven Zusammenhang mit der Leistung in den curriculumsbezogenen Kompetenzen des BiSta-Tests. An dieser Schulart sind reine Mathematikaufgaben häufiger Teil des Mathematikunterrichts als an integrierten Gesamtschulen. Da reine Mathematikaufgaben eher in den curriculumsbezogenen Leistungstests abgefragt werden, findet sich vermutlich eher bei dem BiSta-Test ein positiver Zusammenhang mit schülerorientierten Unterrichtsmethoden. Für integrierte Gesamtschulen, welche einen stärkeren Fokus auf Modellierungsfähigkeiten haben, ist der Zusammenhang zwischen schülerorientierten Unterrichtsmethoden und der Leistung hingegen bei dem PISA-Test positiv. Unter Modellierungsfähigkeiten sind bspw. Fähigkeiten zu verstehen, mathematisch zu argumentieren oder Alltagssituationen in mathematische Modelle zu übersetzen. Fähigkeiten, die eher in PISA getestet werden. Schülerorientierter Unterricht ist damit nicht leistungshemmend oder per se leistungssteigernd, sondern hat dort einen positiven Einfluss, wo der Schwerpunkt des Unterrichts liegt. Daraus leitet sich die These ab, dass schülerorientierte Unterrichtsmethoden erfolgreich sind, wenn Lehrende und Lernende ausreichend Erfahrung damit haben und auch die Art der Inhalte häufig Teil des Unterrichts ist.

Grundsätzlich positiv zu bewerten ist, dass schülerorientierte Unterrichtsmethoden in denjenigen Schularten häufiger im Unterricht eingesetzt werden, in welchen die Lernenden zumindest in einzelnen Fähigkeitsbereichen positiv von schülerorientierten Unterrichtsmethoden profitieren, also an integrierten Gesamtschulen und Schulen mit mehreren Bildungsgängen. Seltener werden diese Methoden eingesetzt in Schularten, bei denen kein positiver Zusammenhang mit der Leistung besteht. Ungeklärt bleibt jedoch, inwieweit sich die positiven wie auch negativen Zusammenhänge von schülerorientierten Unterrichtsmethoden mit der Leistung innerhalb einer Schulart in der Summe gegenseitig aufheben. Formatives Assessment hängt entgegen der Erwartungen negativ mit der Leistung der Lernenden zusammen, und das in allen Schularten. Eine mögliche Erklärung ist, dass formatives Assessment häufiger bei leistungsschwachen Lernenden bzw. Klassen eingesetzt wird. Somit könnte ein kompensierender Effekt gegeben sein, aber trotzdem bleiben die Leistungen der leistungsschwächeren Lernenden geringer als bei leistungsstärkeren Lernenden. Diese Hypothese könnte in zukünftigen Publikationen über Matching-Verfahren getestet werden.

Die Bedeutung der Schulartunterschiede soll an dieser Stelle abschließend zusammengefasst werden. Unterschiede zwischen den Schularten finden sich vor allem in der Schülerorientierung und in den kognitiven Prozessen, Inhalten und Zielen des Mathematikunterrichts und sie stimmen durchaus mit den traditionellen Profilen der Schularten überein. Somit ergeben sich für die Lernenden unterschiedliche Erfahrungsräume im Unterricht. Die jeweils dominierenden Unterrichtsprozesse hängen jedoch innerhalb der Schularten nur sehr eingeschränkt positiv mit den Leistungen der Lernenden zusammen. Weshalb zu hinterfragen ist, ob – zumindest bezogen auf die untersuchten Merkmale – die Verteilung von Lernenden auf die verschiedenen Schularten über das Argument der optimalen Förderung im mehrgliedrigen Bildungssystem gerechtfertigt ist.

Bedeutend sind die Ergebnisse dieser Arbeit auch für die Schul- und Unterrichtsforschung. So konnte gezeigt werden, dass die Analysen der hier vorliegenden Daten zu unterschiedlichen inhaltlichen Ergebnissen führen, je nachdem, ob Schularten berücksichtigt werden oder nicht. Zudem unterscheiden sich die Zusammenhänge zwischen den Schularten teilweise. Für die Analyse und Interpretation von quantitativen Daten in der Schul- und Unterrichtsforschung wäre es daher wünschenswert, die Schularten grundsätzlich als Kon-

trollvariablen einzubeziehen, auch wenn sich die Fragestellung nicht auf Unterschiede zwischen Schularten bezieht.

8.2 Vergleichbarkeit von Fragebogendaten: Implikationen für die Schul- und Unterrichtsforschung

Eine Besonderheit dieser Arbeit ist, dass geprüft wird, inwieweit die Daten geeignet sind, um Vergleiche zwischen den Schularten anzustellen. Ausgehend von den systembedingten Unterschieden zwischen den Schularten, in der Gruppe der Lernenden und der Lehrenden, sowie in den Bildungszielen und -inhalten der Schulen, wurde die These aufgestellt, dass die Erhebung der Konstrukte über die eingesetzten Items zwischen den Schularten nicht per se vergleichbar ist. Zudem wurde die These aufgestellt, dass die Antworten der Lernenden bei Likert-Skalen systematisch zwischen den Schularten verzerrt sind, weil die Tendenzen, Fragen zuzustimmen (ARS) oder mittlere Antwortkategorien zu wählen (MPR), steigen, je geringer die kognitiven Fähigkeiten der Befragten sind (Roßmann 2017). Da sich die Lernenden der verschiedenen Schularten in ihren mittleren kognitiven Fähigkeiten unterscheiden, ist zu vermuten, dass die Antworten in den verschiedenen Schularten unterschiedlich stark durch Antworttendenzen wie ARS und MPR belastet sind. Diesen Thesen folgend wurde in der hier vorliegenden Arbeit geprüft, ob die latenten Konstrukte äquivalent zwischen den Gruppen erfasst werden und somit die Ergebnisse zuverlässig interpretiert werden können. Die vorliegende Arbeit ging jedoch über das rein methodische Vorgehen hinaus und ergründete inhaltlich, worin sich die Schularten unterscheiden und worin mögliche Ursachen mangelnder Vergleichbarkeit liegen. Zusätzlich wurde hinterfragt, welche Bedeutung mögliche Antworttendenzen für die Ergebnisinterpretation haben.

Die Invarianzanalysen zeigen, dass entsprechend der Annahmen für einzelne Konstrukte keine konfigurale Invarianz vorliegt, und somit die Items die Konstrukte zwischen den Schularten nicht vergleichbar erfassen. Bei den Lernenden liegen schulartspezifische Interpretationen für die Einstellung gegenüber der Schule und dem Lernen vor. Bei den Mathematiklehrkräften sind die Überzeugungen zum Unterrichtsziel, Interesse für Mathematik zu wecken, konfigural invariant und somit nicht vergleichbar zwischen den Schularten. Auch die Items, welche das Potential zur kognitiven Aktivierung durch die Variation von Aufgaben messen, scheinen zwischen den Schularten unterschiedlich interpre-

tiert zu werden. Die inhaltlichen Analysen der Items zeigen potentielle Ursachen für die abweichenden Interpretationen.

Für die beiden invarianten Konstrukte bei den Lehrenden – Überzeugungen zum Unterrichtsziel, Interesse für Mathematik zu wecken, und Potential zur kognitiven Aktivierung durch die Variation von Aufgaben – sind die unterschiedlichen Interpretationen vermutlich auf die Itemformulierungen zurückzuführen, welche teils mehrdeutig sind und viel Interpretationsspielraum bieten. Die inhaltliche Analyse verweist darauf, dass die abgefragten Methoden und Ziele zwischen den Schularten teils verschieden interpretiert werden. Beispielsweise scheint sich die Bedeutungszuschreibung zum Kopfrechnen zwischen den Schularten zu unterscheiden, indem Kopfrechnen teils mit mathematischen Routinen und teils mit der Anwendung im Alltag assoziiert wird. Dies könnte unter anderem auf Unterschiede in der Ausbildung und Unterschiede in den Bildungszielen der Schularten zurückgeführt werden.

Bei den Einstellungen der Lernenden gegenüber der Schule und dem Unterricht sind die schulartspezifischen Konstruktrepräsentationen auf die unterschiedlichen Zugangschancen in die Berufs- und Hochschulbildung zwischen den Schularten zurückzuführen. Die eingesetzten Items beziehen sich direkt auf die zukünftigen Bildungschancen, bspw. *„Mich in der Schule anzustrengen, wird mir den Zugang zum Studium ermöglichen.“* Somit werden nicht primär Einstellungen, sondern die wahrgenommenen Bildungschancen und Bildungsbarrieren erfasst – im Sinne Bourdieus die intuitiven Bildungsstatistiken (siehe Abschnitt 3.4.1). Der Vergleich dieser Einstellungsskalen zwischen den Schularten würde somit zu Artefakten in der Interpretation führen. In der Einstellung gegenüber dem Lernen weisen Klassen an Gymnasien die höchsten mittleren Ausprägungen auf. Gefolgt von integrierten Gesamtschulen und Realschulen. Klassen an Schulen mit mehreren Bildungsgängen haben hingegen die niedrigsten mittleren Ausprägungen. Diese Rangfolge spiegelt exakt die Verteilung von Abschlüssen der Schularten wider, wie sie mit den Daten der Bildungsstatistiken für das Schuljahr 2015/2016 beschrieben wurden (siehe Abschnitt 2.2).

Alle Konstrukte, welche konfigural invariant sind und bei denen somit vergleichbare Konstruktrepräsentationen vorliegen, sind auch metrisch invariant. Der Umgang mit der Antwortskala ist somit als vergleichbar zwischen den Schularten anzusehen. Lediglich für die Interpretation des Skalenursprungs kann für eine Skala keine Invarianz (skalare Invarianz) zwischen den Schular-

ten angenommen werden (instrumentelle Motivation der Lernenden). Einschränkungen in der Vergleichbarkeit zwischen den Schularten ergeben sich somit zusammenfassend eher aufgrund inhaltlicher Unterschiede.

Zusätzlich werden die Daten, welche mit Likert-Items erhoben wurden, durch Antworttendenzen verzerrt. Antworttendenzen finden sich bei den Lernenden wie auch bei den Lehrenden. Die Ausprägung der Antworttendenzen unterscheidet sich zwischen den Lehrenden der verschiedenen Schularten jedoch nur marginal. Die Lernenden der verschiedenen Schularten hingegen unterschieden sich entsprechend der Annahme, dass mit geringerer kognitiver Fähigkeit die Wahrscheinlichkeit von Antworttendenzen steigt, stärker in ihren Antworttendenzen, besonders in der Tendenz, Items zuzustimmen (ARS). Die Bedeutung von Antworttendenzen für den Vergleich von Mittelwerten zeigt der Vergleich der Mittelwerte der untersuchten Konstrukte, welche mit und ohne Kontrolle von Antworttendenzen berechnet wurden. Bei den Daten der Lehrenden unterscheidet sich die Beurteilung der Schulartunterschiede bei zwei von vier Konstrukten. Bei einem Konstrukt werden die Schulartunterschiede ohne Kontrolle von Antworttendenzen unterschätzt, beim anderen überschätzt. Die Differenzmuster der Mittelwertunterschiede zwischen den Schularten unterscheiden sich jedoch nicht. Bei den Lernenden hingegen finden sich zusätzlich zu den Unterschieden in der Ergebnisinterpretation bei fünf von 16 Konstrukten auch Unterschiede in den Differenzmustern. Die Schulartunterschiede werden somit auch bei den Daten der Lernenden ohne Kontrolle von Antworttendenzen teils unterschätzt und teils überschätzt. Zudem ändern sich die Abstände und teilweise sogar die Rangfolge der Schularten.

Aus den beschriebenen Ergebnissen ergaben sich Konsequenzen für die weiteren Analysen dieser Arbeit. Zum einen wurden Mittelwertvergleiche und Vergleiche von Zusammenhängen zwischen Schularten nur interpretiert, wenn die erforderliche Invarianzstufe erreicht war. Die Invarianzanalysen konnten jedoch nicht aufdecken, dass die Daten durch Antworttendenzen verzerrt sind. Zudem variieren das Ausmaß und die Richtung, in welche die Daten verzerrt sind. Daher wurden alle vergleichenden Analysen unter Kontrolle von Antworttendenzen berechnet. Es ist jedoch auch zu überlegen, welche Bedeutung die Ergebnisse zur Vergleichbarkeit der Daten allgemein für Fragebogenstudien in der Schul- und Unterrichtsforschung hat. Die beschriebenen Unterschiede zwischen den Schularten, welche vermutlich die Vergleichbarkeit der Daten einschränken, können auch die Vergleichbarkeit von Daten anderer Fragebogen-

studien gefährden. Die folgenden Abschnitte gehen daher auf potentielle Gefahrenquellen ein und beleuchten Möglichkeiten, mangelnde Vergleichbarkeit einzudämmen bzw. mit mangelnder Invarianz umzugehen.

Die inhaltliche Analyse der Items von Konstrukten, welche mangelnde konfigurale Invarianz aufweisen, lässt darauf schließen, dass zum einen strukturelle Unterschiede zwischen den Schularten dazu geführt haben, dass einzelne Konstrukte zwischen den Schularten nicht vergleichbar sind. Zum anderen führen die teils vagen Formulierungen von Items vermutlich dazu, dass einzelne Items je nach Schulart unterschiedlich interpretiert werden. Diese Formen von Konstruktbias und Itembias könnten bereits bei der Entwicklung von Fragebögen adressiert werden (van de Vijver und Leung 2011). In der ländervergleichenden Forschung gibt es bereits Empfehlungen, wie diese Formen von Bias reduziert werden können (bspw. He und van de Vijver 2012), welche auch auf Bias aufgrund von systembedingten Unterschieden zwischen Schularten oder Bildungssystemen auf nationaler Ebene übertragen werden können. In Anlehnung an He und van de Vijver (2012) könnten die Instrumente bspw. simultan für die verschiedenen Schularten entwickelt werden und nur Items, welche für alle Schularten bedeutsam sind, in den Fragebogen aufgenommen werden (decentering). Zusätzlich kann die Interpretation der Items geprüft werden, in dem eine kleine Stichprobe von Lernenden und Lehrenden verschiedener Schularten die Items bearbeiten und kommentieren (Think aloud method; siehe van Someren et al. 1994) und die Interpretation speziell auf mögliche Unterschiede zwischen Schularten geprüft wird.

Trotz sorgfältiger Fragebogenkonstruktion werden eventuell Einschränkungen in der Vergleichbarkeit durch verschiedene Formen von Bias im Vorfeld nicht aufgedeckt, weshalb die Invarianz der Konstrukte beeinträchtigt sein kann. Besonders bei Fragebogenstudien mit internationaler Ausrichtung können nationale Eigenheiten und systematische Unterschiede in der Interpretation zwischen Schularten nur eingeschränkt berücksichtigt bzw. aufgedeckt werden. Antworttendenzen schränken, wie in dieser Arbeit gezeigt wurde, zusätzlich die Vergleichbarkeit der Daten ein. Daher ist es für Fragebogenstudien in der Schul- und Unterrichtsforschung empfehlenswert, generell die Konstrukte auf Invarianz zwischen den Schularten zu prüfen und die Daten nur entsprechend der erreichten Invarianzstufe zu analysieren. Zusätzlich sollte untersucht werden, ob die Daten durch Antworttendenzen verzerrt sind und entsprechend

dafür zu kontrollieren. Mögliche Methoden hierfür wurden in dieser Arbeit vorgestellt.

Für die Interpretation von fragebogenbasierten Ergebnissen der Schul- und Unterrichtsforschung sind die hier aufgeführten Ergebnisse zur Vergleichbarkeit der Daten nicht unproblematisch, soweit Antworttendenzen nicht berücksichtigt und die Invarianz der Messinstrumente zwischen den Schularten nicht geprüft wurden. Zum einen ist unsicher, ob inhaltlich tatsächlich das Gleiche in Bezug gesetzt wird und nicht sprichwörtlich Äpfel mit Birnen verglichen werden. Da zudem keine systematischen Verzerrungen durch Antworttendenzen vorliegen, sondern Differenzen zwischen den Schularten sowohl stärker als auch schwächer werden können und die Mittelwerte in positiver wie negativer Richtung abweichen, ist unklar, wie stark und in welcher Weise die Antworten bei Likert-Skalen durch Antworttendenzen verzerrt sind. Die Studien, welche in dieser Arbeit im Forschungsüberblick dargestellt wurden, scheinen die Vergleichbarkeit der Daten bislang nicht thematisiert zu haben. Diese Ergebnisse gilt es daher mit Vorbehalt und wissend der möglichen Einschränkungen zu interpretieren, da – wie in dieser Arbeit beispielhaft gezeigt wurde – Differenzen möglicherweise auf Artefakte zurückzuführen sind und die Differenzmuster von Schulartunterschieden möglicherweise verzerrt sind. Um die Ergebnislage der vergleichenden Schul- und Unterrichtsforschung abzusichern, wäre es wünschenswert, die verwendeten Daten nachträglich in Sekundäranalysen hinsichtlich der Vergleichbarkeit zu prüfen.

8.3 Grenzen des empirischen Ansatzes der Arbeit und inhaltliche Einschränkungen

Unter Zuhilfenahme quantitativer Methoden wurde im empirischen Teil der Arbeit (Kapitel 7) der Mathematikunterricht von Lernenden in der neunten Klasse vergleichend zwischen den Schularten betrachtet. Hierbei wurden zunächst Invarianzanalysen durchgeführt, um zu prüfen, inwiefern die einzelnen Konstrukte zur Beschreibung der Unterrichtsprozesse, der Unterrichtsziele und der Schülergruppe vergleichbar sind zwischen den Schularten. Zudem wurde geprüft, ob Antworttendenzen die Daten systematisch verzerren. Für die Beschreibung der Schularten (Abschnitt 7.2) wurden die Mittelwerte einzelner Merkmale kovarianzanalytisch unter Kontrolle von Antworttendenzen untersucht. Vermittelnde Mechanismen für Schulartunterschiede in Unterrichtspro-

zessmerkmalen (Abschnitt 7.3) sowie Unterschiede in den Zusammenhängen von Unterrichtsprozessmerkmalen und Leistung (Abschnitt 7.4) wurden mittels Mediations- und Mehrebenen-Moderationsmodellen analysiert. In den folgenden Abschnitten wird reflektiert, welche Limitationen sich durch die eingesetzten Methoden und die verwendete Datengrundlage ergeben. Die Limitationen werden dabei zum einen hinsichtlich der Merkmalsauswahl diskutiert und zum anderen hinsichtlich der Interpretation der Daten. Abschließend wird Bezug zum theoretischen Rahmen der Arbeit genommen.

Die Arbeit hat den Anspruch, den Mathematikunterricht umfassend zwischen den Schularten zu vergleichen. Trotzdem konnten nicht alle im theoretischen Hintergrund hergeleiteten Merkmale vergleichend zwischen den Schularten betrachtet werden. Die Einschränkungen in der Merkmalsauswahl begründen sich durch die Ergebnisse der Invarianzanalyse und durch den für die Analysen verwendeten Datensatz. Die Invarianz der in dieser Arbeit analysierten Konstrukte wurde als Bedingung für den Vergleich der Schularten vorausgesetzt. Dabei wurde eine relativ strenge Methode der Invarianztestung herangezogen, welche in der Forschungspraxis mit der Prüfung der drei Invarianzstufen häufig dazu führt, dass zwar metrische Invarianz erreicht wird, aber keine skalare Invarianz (Robitzsch 2015). Andere Ansätze sind weniger streng in den Annahmen bezüglich metrischer und skalarer Invarianz. Partielle Invarianz erlaubt bspw. dass einige Ladungen oder Intercepts zwischen Gruppen variieren dürfen. Allerdings wird bei diesen Ansätzen die Unsicherheit, die durch die schrittweise Anpassung des Modells hinsichtlich einzelner invarianter Parameter besteht, meist ignoriert (Robitzsch 2015). Für die in dieser Arbeit untersuchten Konstrukte wurde bewusst das strengere Kriterium gewählt, da die Analyse der Schulartunterschiede im Vordergrund stand, welche möglichst gesichert auf tatsächliche Gruppenunterschiede zurückgeführt werden sollen. Ein Ansatz, welcher diese Unsicherheiten bei partieller Invarianz minimiert, ist die Methode des „invariance alignment“ (Muthén und Asparouhov 2014). Auch hier ist es ausreichend, wenn ein Teil der Parameter invariant ist. Es wird jedoch eine Funktion eingeführt, die kleine Abweichungen stärker „bestraft“ als große Abweichungen. Damit findet diese Methode praktisch diejenigen Items, auf denen die Gruppenvergleiche basieren. Mit diesem Ansatz kann somit direkt bei der Skalierung partielle Invarianz analysiert und die Skalenwerte entsprechend berechnet werden. Hierbei werden die Skalenwerte vergleichbar, während die Messinvarianz minimiert wird (Muthén und Asparouhov 2014). In der hier vorliegenden Arbeit handelt es sich jedoch um Sekundäranalysen der

PISA-2012-Skalen, für welche die partielle Invarianz bei der Skalierung nicht berücksichtigt wurde. Dementsprechend muss bei der Invarianzprüfung der Konstrukte auf vollständige Invarianz geprüft werden.

Die in dieser Arbeit durchgeführten Invarianzanalysen gilt es auch im Kontext von Antworttendenzen zu reflektieren. Invarianzanalysen können zwar nicht immer die Verzerrung der Daten durch Antworttendenzen aufdecken (Liu et al. 2017), dennoch können Antworttendenzen als eine Form des Methodenbias die Invarianz von Messinstrumenten reduzieren (siehe bspw. Cheung und Rensvold 2000; Grimm und Church 1999). In den durchgeführten Invarianzanalysen der hier vorliegenden Arbeit wurden Antworttendenzen jedoch nicht berücksichtigt. Daher könnte die Invarianz der hier untersuchten Konstrukte unterschätzt worden sein. Aufbauend auf die hier vorliegende Arbeit könnten in einem komplexeren Analysemodell die Ausgangsdaten korrigiert werden, indem zunächst in einer Regression des jeweiligen Konstrukts auf die Antworttendenzen (ARS-Index und MPR-Index) die gemeinsame Varianz sowie die Residuen berechnet werden. In einem zweiten Schritt würde die Invarianzanalyse mit den Residuen durchgeführt, welche dann die korrigierten Werte sind. Dieses komplexe methodische Vorgehen hätte den Rahmen der hier vorliegenden Arbeit, welche inhaltlich Unterschiede zwischen Schularten untersucht, gesprengt. In einer zukünftigen Publikation mit methodischem Schwerpunkt könnte der Zusammenhang von Antworttendenzen und mangelnde Invarianz jedoch untersucht werden. Es wäre interessant zu vergleichen, ob sich die Ergebnisse der Invarianzanalyse ändern, wenn wie hier beschrieben die Antworttendenzen kontrolliert werden. Eventuell könnten mehr Konstrukte für die Beschreibung der Schularten herangezogen werden, aber immer vorausgesetzt, dass Antworttendenzen auch in den weiteren Analysen kontrolliert werden. Durch das strenge Vorgehen in dieser Arbeit konnten keine affektiven Merkmale in den Vergleich einbezogen werden, da die Datenqualität für mehrere affektive Merkmale nicht ausreichend war. So konnte nur das Interesse der Lernenden für das Fach Mathematik in den Vergleich der Schularten eingehen, welches sich jedoch nicht bedeutend zwischen den Schularten unterscheidet. Affektive Merkmale wären jedoch nicht nur für die Beschreibung der Schülergruppe zentral, sondern auch für die Untersuchung des Einflusses von Unterricht auf eben diese Merkmale.

Zusätzlich zu der strengen Prüfung auf Invarianz wurde die Auswahl der für die Beschreibung verwendeten Merkmale auch durch die Studie selbst beschränkt.

Die hier vorliegende Arbeit basiert auf Sekundäranalysen der PISA-2012-Studie bzw. deren nationaler Erweiterung. In der PISA-Studie werden Merkmale des Unterrichts, der Lehrenden und der Schülergruppe vergleichsweise umfassend erfasst. Trotzdem fehlen Merkmale, die für die Fragestellungen dieser Arbeit von Interesse sind. Dies sind bei den Lehrenden insbesondere das Professionswissen und Überzeugungen bezüglich subjektiver Lerntheorien und subjektive Theorien über das Lehren des Gegenstandes. Bei den Lernenden wären Merkmale über die Interaktion innerhalb der Lerngruppe interessant für die Analyse von Peer-Effekten. Unterrichtsprozesse und Unterrichtsinhalte sind in dem verwendeten Datensatz umfassend erfasst, jedoch auf einer sehr abstrakten Ebene. Die Auswahl und Operationalisierung der Konstrukte in den PISA-Studien findet auf internationaler Ebene statt und muss mit dem Fokus auf den internationalen Vergleich inhaltlich zu den Bildungssystemen der teilnehmenden Länder passen. Dadurch sind tiefergehende Inhalte selten realisierbar, bspw. kann das realisierte Curriculum nur grob erfragt werden, da sich die Curricula zwischen den Länder stark unterscheiden (vgl. Thompson et al. 2017). Bislang scheint es jedoch keinen alternativen Datensatz zu geben, der alle interessierenden Merkmale abdeckt und zudem mit einer angemessenen Stichprobengröße vergleichende Analysen zwischen Schularten auf Klassenebene zulässt.

Neben den Einschränkungen in der Merkmalsauswahl ist die Arbeit auch hinsichtlich der Ergebnisinterpretation eingeschränkt, da Daten für die Schulart Hauptschule sowie Unterschiede zwischen Bundesländern nicht berücksichtigt werden konnten. Bei der ersten Erhebungswelle im Jahr 2012 wurden zwar auch Lernende an Hauptschulen befragt bzw. getestet. Bei der zweiten Erhebungswelle im Jahr 2013 ist jedoch ein großer Stichprobenausfall bei den Lernenden an Hauptschulen zu verzeichnen, wodurch die Zuverlässigkeit der Daten für Lernende an Hauptschulen nicht gewährleistet ist. Ein Grund hierfür ist, dass es an Hauptschulen nur teilweise eine verpflichtende 10. Klasse gibt. Dadurch wird in der relativen Gegenüberstellung der Schularten die Hauptschule nicht berücksichtigt, was die Interpretation der Schulartunterschiede beeinflussen kann. Aufgrund der Fallzahl konnten auch Abweichungen zwischen den Bundesländern nicht berücksichtigt werden. Durch das föderale Bildungssystem in Deutschland variiert das Schulartangebot zwischen den Bundesländern und auch in der inhaltlichen Ausgestaltung der Schularten finden sich organisatorische wie curriculare Abweichungen. Vor allem bei den Schulen mit mehreren Bildungsgängen ist die Diversität der Schulen relativ hoch. Trotzdem kann-

ten für einzelne Merkmale systematische Unterschiede zwischen den Schularten festgestellt werden, vor allem in der inhaltlichen Ausgestaltung des Mathematikunterrichts. Die Einführung der Bildungsstandards trägt vermutlich dazu bei, dass sich die Schularten über die Bundesländer hinweg inhaltlich stärker angleichen. Interessant wäre, die hier durchgeführten Analysen mit einem geeigneten Datensatz gezielt für einzelne Bundesländer durchzuführen. Es ist zu vermuten, dass durch die stärkere Homogenität der Schularten innerhalb der Bundesländer die Differenzen zwischen den Schularten innerhalb eines Bundeslandes größer sind.

Weitere Einschränkungen in der inhaltlichen Interpretation ergeben sich durch die eingesetzten Methoden. Zum einen ist nicht sicher, dass sich immer alle Schularten bedeutend voneinander unterscheiden. Dies liegt an dem verwendeten Maß zur Beurteilung der Mittelwertunterschiede. Das Effektstärkemaß Eta Quadrat gibt an, ob sich mindestens eine Schulart bedeutend von den anderen Schularten unterscheidet. Somit ist nicht sicher, ob sich wirklich immer alle Schularten bedeutend voneinander unterscheiden, auch wenn das Effektstärkemaß Eta Quadrat einen mittleren oder hohen Effekt der Schulart wiedergibt. Zum anderen ist die Interpretation der Moderations- und Mehrebenen-Mediationsanalysen auf den Vergleich des Gymnasiums mit Realschulen, integrierten Gesamtschulen und Schulen mit mehreren Bildungsgängen beschränkt, da die Schularten als Dummy-Variablen in die Modelle aufgenommen wurden mit dem Gymnasium als Referenzkategorie. Die Unterschiede werden also immer zwischen dem Gymnasium und einer der anderen Schularten interpretiert. Ein direkter Vergleich zwischen bspw. Realschulen und integrierten Gesamtschulen erfolgt nicht. Das Gymnasium wurde als Referenzkategorie gewählt, da diese Schulart in den deskriptiven Beschreibungen bei den meisten Merkmalen die größten Abweichungen zu den anderen Schularten gezeigt hat. Um die Unterschiede zwischen den anderen Schularten statistisch abzusichern, müssten die Analysen wiederholt werden, mit jeweils einer der anderen Schularten als Referenzkategorie. Bei den Moderationsanalysen wären alternativ auch Invarianzanalysen denkbar. Dann hätten die Merkmale aufgrund der geringen Fallzahl jedoch nicht auf Klassenebene, sondern auf Individualebene durchgeführt werden müssen. Dies hätte jedoch die inhaltliche Interpretation geändert.

Die Ergebnisse dieser Arbeit werden wiederholt mit den allgemeinen Profilen der Schularten in Bezug gesetzt. Diese wurden aus der historischen Entwicklung der Schularten, wissenschaftlichen Beschreibungen der Schularten und

den amtlichen Statistiken hergeleitet. Bei den amtlichen Statistiken ist es jedoch nicht immer möglich, Schulart und Bildungsgang zu trennen, besonders bei den Abschlüssen und Anschlussperspektiven. Zudem sei darauf verwiesen, dass die verwendete Literatur, welche gezielt einzelne Schularten beschreibt, teilweise zehn Jahre alt ist. Somit ist der Bezug der Ergebnisse zu den Schulartprofilen teilweise eingeschränkt. Inhaltlich wurden die Aussagen über die Schularten in dieser Arbeit jedoch mit den aktuellen Schulartbeschreibungen der Kultusministerien gegengeprüft. Allerdings konnte keine tiefgreifende Analyse der Selbstbeschreibungen durchgeführt werden. In dieser Arbeit wurde gezeigt, dass die ideellen Vorstellungen, die mit den jeweiligen Schularten verbunden sind, sich in den Zielen, Inhalten und Prozessen des Unterrichts widerspiegeln. Es scheint lohnenswert, diese Selbstbeschreibungen zu systematisieren und gegenüberzustellen, um somit ein aktuelles Bild der Schularten festzuhalten. Parallel zur Entstehung dieser Arbeit haben sich zudem in einzelnen Bundesländern (bspw. in Berlin) Änderungen in der Ausgestaltung des Bildungssystems vollzogen. Die Erkenntnisse aus der Begleitforschung konnten noch nicht in den theoretischen Hintergrund dieser Arbeit einfließen. Wie die Schularten konkret gestaltet sind und inwieweit sich dadurch Änderungen für den Unterricht an den Schularten ergeben, bleibt spannend zu verfolgen. Die hier vorliegende Arbeit könnte mit der Beschreibung und Analyse der Schularten im Jahr 2012 eine hilfreiche Ausgangsbasis für die Interpretation der Ergebnisse zu der Schulstruktur und den Schularten in den sich wandelnden Bildungssystemen sein.

8.4 Schulstruktur oder Einzelschule – was zählt?

In der nationalen Bildungspolitik sind Schulstrukturfragen in den Hintergrund getreten, denn es gibt kaum ein bildungspolitisches Thema, das über viele Jahre bei den Parteien so strittig war wie die Frage der Schulstruktur in der Sekundarstufe (Tillmann et al. 2008). Fends abschließendes Statement aus seinen Gesamtschulstudien, dass es mehr auf die Einzelschule als auf die Schulstruktur ankomme, war aus politischer Sicht vermutlich eine dankbare Schlussfolgerung. Die Schule wird seitdem weniger als ein starr fixiertes und weitgehend vereinheitlichtes Gefüge, sondern vielmehr als gestaltbarer Handlungs-, Lebens- und Erfahrungsraum aufgefasst. Begriffe wie Schulqualität oder Schuletikos verweisen auf die Herausforderung aller Beteiligten zur gemeinsamen, verantwortlichen Gestaltung von Schule (Ditton und Kreyer 1995). Im Zuge der PISA-Studie drohte die Debatte um die Schulstruktur wieder aufzuflam-

men. Zumindest die Zahl der Presseartikel, welche die Schulstruktur thematisieren, sind mit der Veröffentlichung der ersten PISA-Studie stark angestiegen (Tillmann et al. 2008). Von politischer Seite aus wurde das Thema hingegen vermieden und nicht auf den Handlungskatalog gesetzt. Stattdessen stand die Frage der Qualitätssicherung im Fokus. War anfangs die Qualität der Schule im Zentrum der Diskussion, hat sich der Fokus inzwischen auf die Qualität des Unterrichts und, nicht zuletzt mit der „Qualitätsoffensive Lehrerbildung“ der KMK, auf die Lehrenden verschoben.

Betrachtet man die Ergebnisse der hier vorliegenden Arbeit unter diesem neuen Blickwinkel, könnte die Schulstruktur und somit die Bedeutung der Schulart als vernachlässigbar angesehen werden. Insgesamt sind die Unterschiede zwischen den Schularten geringer als angenommen. Mit Bezug zu Baumerts „differenziellen Lern- und Entwicklungsmilieus“ (Baumert et al. 2006b) wurde angenommen, dass sich die Schularten hinsichtlich der affektiven Merkmale der Lernenden unterscheiden. Die empirischen Befunde Baumerts und Kollegen (2006) können in dieser Arbeit jedoch nicht repliziert werden. Auch bezüglich der Überzeugungen von Lehrenden und der Unterrichtsprozessmerkmale finden sich entgegen der Erwartung nur wenige bedeutende Schulartunterschiede. Es ist daher zu fragen, ob Schulartunterschiede in den deskriptiven Vergleichen anderer Studien bislang überschätzt wurden und, ähnlich wie es Fend für seine vergleichenden Analysen von Gesamtschulsystem und traditionellem Schulsystem schlussfolgert, der Einzelschule die größere Bedeutung zukommt (Fend 1982, 1988).

Diese Frage ist mit einem klaren „Ja“ zu beantworten. Für manche Merkmale werden Schulartunterschiede durch den rein deskriptiven Vergleich von Mittelwerten vermutlich überschätzt. Dies sei am Beispiel des Merkmals Disziplin verdeutlicht, welches in den Analysen in dieser Arbeit nach den Leistungsmerkmalen und den Merkmalen sozioökonomischer Herkunft die größten Mittelwertdifferenzen aufweist. Bezieht man jedoch mit ein, dass dieses Merkmal innerhalb der Schularten stark variiert, sind die Mittelwertunterschiede zwischen den Schularten unbedeutend. Merkmale jedoch, welche direkt in den Beschreibungen der Schularten als Alleinstellungsmerkmale betont werden (siehe Kapitel 2), zeigen die erwarteten Unterschiede, auch wenn die Varianz des Merkmals innerhalb der Schularten berücksichtigt wird. Hervorzuheben sind hier Merkmale, welche Unterrichtsinhalte und schülerorientierte Unterrichtsmethoden wie das formative Assessment beschreiben oder auch Überzeu-

gungen von Lernenden, welche sich auf kognitive Ziele des Mathematikunterrichts beziehen.

Der Schulstruktur und der Schule kommt demnach je nach untersuchtem Merkmal unterschiedliche Bedeutung zu. Einflüsse durch das gegliederte Bildungssystem sind vor allem dann zu erwarten, wenn die Merkmale in direktem Zusammenhang mit intendierten Schulartunterschieden stehen. Dieser Abschnitt endet daher mit einem Zitat, welches ebenfalls Fends Schlussfolgerungen zu den Gesamtschulstudien entnommen ist, das jedoch nicht so prominent verbreitet wurde wie seine Bedeutungszuschreibung zur Einzelschule: „Es sei an dieser Stelle aber betont, daß es verfehlt wäre, die Systemfrage gegen die Qualität der einzelnen Schulen in gegenseitiger Abschätzung auszuspielen. Es dürfte sich mit hinreichender Deutlichkeit gezeigt haben, daß die strukturellen Vorgaben eine Bedeutung haben [...]“ (Fend 1982, S. 501).

8.5 Schularten im Kontext gesellschaftlicher, wirtschaftlicher und individueller Anforderungen und Entwicklungen

In der vorliegenden Arbeit wurden schwerpunktmäßig Unterschiede zwischen den Schularten der Sekundarstufe I herausgearbeitet. Hierbei wurde auch die Passung der empirischen Befunde zum Mathematikunterricht mit den allgemeinen Beschreibungen der Schularten in Bezug gesetzt. In diesem Abschnitt soll abschließend darüber hinaus die Bedeutung und Relevanz der Schularten hinsichtlich allgemeiner gesellschaftlicher und bildungspolitischer Entwicklungen diskutiert werden. Verschiede Schularten mit je eigenen Schwerpunkten bieten das Potential, gezielt auf unterschiedliche Lebensläufe vorzubereiten und den Bedürfnissen von Seiten der Lernenden gerecht zu werden. Die Frage ist nun, ob die Schularten in ihrer aktuellen Ausgestaltung diesen Auftrag erfüllen können und den individuellen Bedürfnissen und Zielen sowie den wirtschaftlichen Entwicklungen gerecht werden.

Zunächst wird die Hauptschule getrennt von den anderen Schularten beleuchtet. Diese Schulart konnte im empirischen Teil dieser Arbeit nicht berücksichtigt werden, weshalb sich die Diskussion bei dieser Schulart auf das herausgearbeitete Schulartprofil (Kapitel 2) stützt sowie auf den Forschungsstand (Ab-

schnitt 3.5) bezieht. Der Schwerpunkt der Hauptschule liegt in der Berufsvorbereitung. Zudem kennzeichnet die starke Bindung von Lernenden und Lehrenden durch das Klassenlehrerprinzip das Profil dieser Schulart. Der Unterricht zeichnet sich durch eine vergleichsweise hohe Schülerorientierung bzw. konstruktive Unterstützung aus (Baumert et al. 2004; Baumert und Kunter 2006; Bayer et al. 2015; Gruehn 2000; Klieme und Rakoczy 2001; Klieme, Schümer et al. 2001; Tillmann und Meier 2001). Dieses Profil wird jedoch von Seiten der Lernenden bzw. deren Eltern nicht mehr gut angenommen, was zum einen an den geringen Anschlussperspektiven liegen mag, welche die Hauptschule bietet. Beispielsweise sind nur 40 % der Lehrstellen für Jugendliche mit Hauptschulabschluss offen (Anbuhl 2016). Auch der Übergang an Schularten, welche einen höheren Abschluss bieten, ist selten (siehe Abschnitt 2.2.1). Somit scheint die Hauptschule den zunehmenden Bildungsaspirationen als auch den Veränderungen in der subjektiven Bewertung von Kosten und Nutzen für eine höhere Bildung nicht mehr gerecht werden. Der Schwerpunkt in der Schülerorientierung scheint auch nicht zur Wahl dieser Schulart zu motivieren, wenn den Lernenden eine alternative Schulform im allgemeinen Schulsystem offen steht. Oben genannte Zahlen zum Lehrstellenangebot verdeutlichen zum anderen, dass von Seiten der Arbeitgeber der Hauptschule eine fundierte Berufsvorbereitung abgesprochen wird. Trotz der praxisnahen Vorbereitung und der Kooperation mit Unternehmen werden für die meisten Ausbildungsberufe höhere Abschlüsse als ein Hauptschulabschluss erwartet. Inzwischen kann an Hauptschulen zwar meist über einen verlängerten Schulbesuch auch der mittlere Abschluss erworben werden, was im Schuljahr 2015/2016 34 % der Lernenden realisiert haben (Statistisches Bundesamt, GENESIS-Online Datenbank). Interessant ist jedoch, ob von Seiten der Betriebe dieser Abschluss gleichwertig zu einem mittleren Abschluss behandelt wird, der an einer anderen Schulart erworben wurde. Der sehr hohe Anteil an Hauptschülern, die in das Übergangssystem wechseln (bspw. in ein Berufsvorbereitungsjahr), spricht nicht dafür. Somit scheint durch die wachsenden Bildungsansprüche und wachsende Anforderungen in den Ausbildungsberufen die Hauptschule keine attraktive Schulform zu sein und den Bildungszielen nicht mehr gerecht zu werden. Die wörtliche Rolle der „Hauptschule“ hat diese Schulart längst verloren und als Schule der „Praktisch Begabten“ scheint sie auch ausgedient zu haben. Spätestens mit der Diskussion in den Medien um die „Hauptschule als Chance für geflüchtete Kinder“ (Trautsch 2016) wird klar, dass die Hauptschule nicht mehr mit dem Ziel von berufsvorbereitender Qualifikation verbunden wird, sondern vielmehr mit der Vermittlung der deutschen Sprache und Akkulturation.

Damit kann festgehalten werden, dass die Hauptschulen ihre eigentlichen Bildungsziele nicht mehr erfüllen (können). Trotzdem besuchen Lernende noch Hauptschulen, wenn auch mit sinkender Tendenz. Was bedeutet es für diese Lernenden, wenn Hauptschulen verschwinden? Es gibt durchaus Befürworter der Hauptschule, welche die integrative Funktion, die individuellen Fördermöglichkeiten durch kleine Klassen, die hohe Schülerorientierung sowie die Expertise im Umgang mit ethnisch gemischten Klassen und in Deutsch als Fremdsprache hervorheben. Im Umkehrschluss wird befürchtet, dass Abbrecherquoten steigen, wenn Lernende an anderen Schularten unterrichtet würden (bspw. Schavan et al. 2008). Diesen Argumenten ist entgegenzusetzen, dass die Expertise nicht verschwindet, wenn es keine Hauptschulen mehr gibt. Vielmehr besteht das Potential, die Expertise an anderen Schularten einzubringen, wodurch auch die Lernenden anderer Schularten von der Expertise profitieren. Kleinere Klassen oder niedrigere Betreuungsschlüssel sollten nicht nur an der Hauptschule, sondern an allen Schularten realisiert werden. Vor allem vor dem Hintergrund der steigenden Individualisierung des Lernens innerhalb des Klassenverbandes und einer Verbesserung der Inklusion sind kleinere Klassen an allen Schularten wünschenswert. Berufsorientierung, praktische Angebote und eine Orientierung am Lernenden gibt es auch in Schulen mit mehreren Bildungsgängen, integrierten Gesamtschulen und Realschulen. Wobei die Anschlussperspektiven an diesen Schularten besser sind als an Hauptschulen und bei den angestrebten Abschlüssen variabler auf die Entwicklung der Lernenden eingegangen werden kann. Inwieweit diese Angebote die Bedürfnisse von Lernenden, welche eine Hauptschule besuchen, in der Praxis adressieren, kann an dieser Stelle nicht beantwortet werden. Bislang scheint die Abbrecherquote in Bundesländern, welche die Hauptschule abgeschafft haben, nicht nachweislich gestiegen. Somit bleibt zu überlegen, ob Hauptschulen weiterhin Bestand haben sollten oder integrativere Formen eine sinnvolle Alternative sind.

Auch eigenständige Realschulen sind auf dem Rückzug, obwohl diese deutlich stärker nachgefragt waren und sind als Hauptschulen. Ein Grund für den sinkenden Anteil an Realschulen ist unter anderem, dass bei der Umwandlung von Hauptschulen in stärker integrierte Formen teilweise Haupt- und Realschulen zusammengelegt wurden und teils ganze Bundesländer Haupt- und Realschulen durch mehrgliedrige Schularten ersetzt haben (Autorengruppe Bildungsberichterstattung 2016). Im Schuljahr 2012/2013, auf welches sich auch die in dieser Arbeit verwendeten Daten beziehen, waren Realschulen noch in 9 von 16 Bundesländern vertreten, mit sinkender Tendenz (Statistisches Bundesamt,

GENESIS-Online Datenbank). In der Mehrheit der Bundesländer stand die Realschule zur Zeit der Datenerhebung damit in direkter Konkurrenz zu den anderen Schularten, weshalb die hier vorliegenden Analysen gut geeignet sind, die verschiedenen Schularten inhaltlich gegenüberzustellen und im Kontext der gesellschaftlichen und wirtschaftlichen Entwicklungen zu reflektieren.

Realschulen haben neben den Hauptschulen den stärksten Fokus in der Berufsvorbereitung. Dies schlägt sich besonders im Fächerkanon nieder, welcher gezielt auf Berufe entweder in wirtschaftlichen, sprachlichen oder mathematisch-technisch-naturwissenschaftlichen Bereichen vorbereitet. Integrierte Gesamtschulen und Schulen mit mehreren Bildungsgängen zielen zwar auch auf die Berufsvorbereitung, bilden aber weniger gezielt auf bestimmte Arbeitsbereiche hin aus. Teilweise gibt es jedoch schulspezifische Schwerpunktlegungen. Auf den Mathematikunterricht bezogen zeigen die in dieser Arbeit durchgeführten Analysen, dass Realschulen in den Inhalten und Zielen des Unterrichts näher an den Gymnasien sind als integrierte Gesamtschulen und Schulen mit mehreren Bildungsgängen. So werden an Realschulen häufiger reine Mathematik und mathematische Konzepte unterrichtet. Zudem steht das Ziel der Mathematisierungsfähigkeit stärker im Fokus als an integrierten Gesamtschulen und Schulen mit mehreren Bildungsgängen.

Trotz der inhaltlichen Unterschiede kann an allen drei Schularten der mittlere Abschluss erworben werden, welches auch der Abschluss ist, den die größte Anzahl von Lernenden an diesen Schularten erwirbt (siehe Abschnitt 2.2). Damit stammt der größte Anteil von Lernenden, welche in die Berufsausbildung einmünden, von einer dieser drei Schularten. Trotz des inhaltlichen Schwerpunktes dieser drei Schularten in der Berufsvorbereitung scheinen die Lernenden nicht ausreichend auf die Anforderungen in den Ausbildungsberufen vorbereitet zu werden. Laut einer Umfrage der IHK sinkt der Anteil der Ausbildungsbetriebe, die zufrieden mit der Ausbildungsreife der Ausbildungsplatzbewerber sind, stetig und liegt 2017 unter zehn Prozent. Von Seiten der Unternehmen werden vorwiegend fehlende mathematische und sprachliche Grundfähigkeiten sowie mangelnde soziale Kompetenzen beklagt (DIHK 2017). Auf der anderen Seite sind die Anforderungen in den Ausbildungsberufen in den letzten Jahren gestiegen. Von den Fachkräften werden nach einer Analyse der Rahmenpläne von verschiedenen Ausbildungsberufen zunehmend Fremdsprachenkenntnisse, berufliche Handlungskompetenz (bspw. fachübergreifendes Wissen, selbständiges Arbeiten und lernen, sowie Reflexion des Handelns) und

soziale Kompetenzen erwartet. Zudem sind die Lernanforderungen kognitiv komplexer. Theoretisches Vorwissen verliert im Gegenzug an Bedeutung (vgl. Protsch 2013). Somit scheint zum einen die Vermittlung von Grundlagen in sprachlichen und mathematischen Fähigkeiten an diesen Schularten für Lernende mit mittlerem Abschluss nicht gesichert. Auf der anderen Seite scheint der Fokus auf beruflicher Qualifizierung in speziellen Aufgabenbereichen nicht mehr passend zum Arbeitsmarkt mit zunehmend flexibleren Anforderungen an die Bewerber. Zunehmend attraktiv für den Ausbildungsmarkt sind Lernende mit Abitur, welche – entsprechend der Verteilung der Lernenden auf die Schularten – zum größten Teil an Gymnasien unterrichtet wurden. Natürlich muss hierbei berücksichtigt werden, dass die Zahl der Abiturienten gestiegen ist. Auf der anderen Seite ist es entgegen der Erwartungen, dass Lernende an Gymnasien steigenden Zuspruch erhalten, obwohl sie nicht gezielt auf die praktischen Tätigkeiten gängiger Ausbildungsberufe hin beschult wurden. Vielleicht sind Lernende an Gymnasien aber auch gerade, weil an dieser Schulart selbstständiges Denken, Problemlöse- und Handlungskompetenz im Fokus steht, zunehmend qualifiziert für die komplexeren und flexibler werdenden Anforderungen der beruflichen Ausbildung.

Damit Lernende an integrierten Gesamtschulen, Schulen mit mehreren Bildungsgängen und Realschulen weiterhin attraktiv für den Ausbildungsmarkt bleiben, wäre es wünschenswert, zum einen die Vermittlung von Grundfähigkeiten in Mathematik und den Sprachen zu verstärken. Dies betrifft besonders Schulen mit mehreren Bildungsgängen und integrierte Gesamtschulen, welche in mathematischen und sprachlichen Leistungstests im Mittel schlechter abschneiden als Lernende an Realschulen und auch hinsichtlich der Inhalte des Mathematikunterrichts – wie die Ergebnisse der hier vorliegenden Arbeit andeuten – seltener reine Mathematik und mathematische Konzepte behandeln. Zum anderen scheint es in Anbetracht der komplexeren und flexibleren Anforderungen überlegenswert, statt bereichsspezifischer Berufsvorbereitung stärker übergreifende Kompetenzen zu fördern, wie selbständiges Handeln, Kommunikation und Präsentation sowie Reflexionsvermögen. Vor allem in Zuge der Umgestaltung von Bildungssystemen, wie sie aktuell in einzelnen Bundesländern vollzogen wird, bietet sich die Möglichkeit, die dargelegten Schwachstellen anzugehen und zukunftsprospektiv die sich wandelnden Anforderungen des Arbeitsmarktes zu adressieren.

Mit der Forderung nach weniger Spezialisierung und höherer Flexibilisierung wird gleichzeitig die Frage aufgeworfen, ob es die Vielfalt an Schularten braucht. Die intendierten Unterschiede bezüglich der unterrichteten Inhalte scheinen aus heutiger Perspektive weniger gerechtfertigt. Von Seiten des Arbeitsmarktes sind weniger Spezialisierung als viel mehr übergreifende Fähigkeiten gefragt. Von gesellschaftlicher wie auch politischer Seite aus werden zunehmend flexiblere Bildungsmöglichkeiten der Lernenden gewünscht bzw. ermöglicht. Somit müssten auch die Bildungsinhalte zwischen den Schularten einheitlicher werden, um Übergänge zwischen den Schularten und Bildungsgängen zu erleichtern sowie flexibel auf eine berufliche Ausbildung als auch die universitäre Weiterbildung vorbereitet zu sein. Anstatt also kaum unterscheidbare Kompetenzziele für einzelne Fächer in den Curricula zu formulieren (siehe Abschnitt 3.3.2), welche dann wiederum entsprechend der schulartspezifischen Bildungsziele mit unterschiedlichen Inhalten gefüllt werden, könnte die Diversität der Schularten reduziert werden; vor allem da, wie gezeigt wurde, inhaltliche Passungsprobleme zwischen den Schularten und den gesellschaftlichen und wirtschaftlichen Anforderungen bestehen. Wie ein solches Bildungssystem konkret gestaltet sein könnte, kann aus den Erkenntnissen der hier vorliegenden Arbeit nicht abgeleitet werden. Die aktuelle Begleitforschung zu Schulstrukturreformen in einzelnen Bundesländern könnte hingegen Erkenntnisse liefern, welche Konsequenzen für Lehrende und Lernende resultieren, wenn die Schulstruktur auf ein zweigliedriges System reduziert wird. Von besonderem Interesse ist, ob sich die Erfolgs- bzw. Abbrecherquote ändern, wie sich die Lernenden auf die Schularten verteilen, wie gut die Lernenden auf die Ausbildung bzw. das Studium vorbereitet werden, ob es schwierig ist für Lehrende aus traditionellen Schularten in den mehrgliedrigen Schularten zu unterrichten bzw. wie Lehrende auf die neuen Aufgaben vorbereitet werden können.

Interessanterweise ist in keinem Bundesland eine Tendenz zum Einheitsschulsystem erkennbar. Das Gymnasium ist die einzige Schulart, welche im Zuge von Reformen unangetastet bleibt. Für diese Schulart hat sich in der Vergangenheit gezeigt, dass hier das größte Konfliktpotential besteht, wenn die Sonderstellung bedroht scheint. Zweigliedrige Schulsysteme mit dem Gymnasium als eigenständige Schulart sind von politischer Seite vermutlich ein Weg, um eine mögliche Eskalation der Schulstrukturfrage wie in den 1970er Jahren zu unterbinden. Auf der anderen Seite könnten mit dem Angebot von bspw. zwei Schularten die individuellen Bedürfnisse der Lernenden evtl. besser adressiert werden. Die in dieser Arbeit durchgeführten Analysen legen die Vermutung

nahe, dass eine hohe Schülerorientierung in Klassen an integrierten Gesamtschulen und Schulen mit mehreren Bildungsgängen positiv mit der Leistungsentwicklung zusammenhängt, wohingegen die Zusammenhänge in Realschulen und Gymnasien negativ sind. Es ist zu vermuten, dass entsprechend der individuellen Bedürfnisse gezielt die Schulart gewählt wird, die in ihrem ausgeschriebenen Profil am besten zu den Bedürfnissen der Lernenden passt. Dies würde auch mit Forschung zu Übergangsentscheidungen im zweigliedrigen Bildungssystem übereinstimmen. Hier konnte Klinge (2016) vier Typen von Eltern identifizieren, welche sich in der Wahl der Schulart und den zugrundeliegenden Argumentationen unterscheiden. Leistungsförderung, Leistungsdruck, Wohlfühlen, offene Bildungswege, soziale Kompetenzen sowie Berufsvorbereitung sind Themenbereiche, welche Eltern als entscheidungsrelevant ansehen und eher zur Wahl eines Gymnasium oder einer integrierten Schulform motivieren. Vielleicht können gerade die Gegensätze von zwei Schularten wie dem Gymnasium und der integrierten Gesamtschule die unterschiedlichen Bedürfnisse adressieren.

Auch wenn aus der hier vorliegende Arbeit keine direkten Erkenntnisse für die Ausgestaltung der Schulstruktur gefolgert werden können, bietet sie dennoch Einblicke in die Ausgestaltung des Unterrichts der Schularten. Diese können als Verstehensgrundlage zum einen für die Interpretation von Forschungsergebnissen beitragen und zum anderen zum Verstehensprozess der Schularten selbst beitragen. Um verlässliche Aussagen über die zukünftige Ausgestaltung der Schulstruktur zu generieren, sind zusätzliche Forschungen nötig, die verschiedene Perspektiven und Nutzer berücksichtigen. Neben den Wünschen der Eltern und Bedürfnissen der Lernenden sind auch die Erfordernisse in Arbeitsmarkt und Hochschule einzubeziehen. Wertvolle Erkenntnisse lassen sich auch über den Vergleich mit ähnlichen Bildungssystemen in anderen Ländern oder der Ausgestaltung von integrierten Bildungssystemen gewinnen. Besonders aber werden die Erkenntnisse aus den Begleituntersuchungen aktueller Reformbestrebungen hin zum zweigliedrigen Bildungssystem bedeutend sein. Die hier vorliegende Arbeit bietet mit der Bestandsaufnahme aus 2012 eine Vergleichsgrundlage und Beispiele, welche Merkmale über den reinen Vergleich von Leistung und Komposition hinaus interessante Erkenntnisse liefern, besonders über das Kerngeschäft von Schule, den Unterricht.

9 Literaturverzeichnis

- Abdi, H. (2004). Factor rotations in factor analysis. In M. Lewis-Beck, A. Bryman & T. Futing Liao (Hrsg.), *The SAGE encyclopedia of social science research methods*. Thousand Oaks, California: Sage Publications, Inc.
- Abel, W. (1981). Massenarmut und Hungerkrisen in Deutschland im letzten Drittel des 18. Jahrhunderts. In U. Herrmann (Hrsg.), *Das pädagogische Jahrhundert. Volksaufklärung und Erziehung zur Armut im 18. Jahrhundert in Deutschland* (Geschichte des Erziehungs- und Bildungswesens in Deutschland, S. 29–52). Weinheim: Beltz.
- Abs, H. J. (2006). Zur Bildung diagnostischer Kompetenz in der zweiten Phase der Lehrerbildung. *Zeitschrift für Pädagogik* 51, 217–234.
- Adick, C. (2005). Transnationalisierung als Herausforderung für die International und Interkulturell Vergleichende Erziehungswissenschaft. *Tertium comparationis* 11 (2), 243–269.
- Adler, P. A. & Adler, P. (2003). *Peer power. Preadolescent culture and identity* (3. Aufl.). New Brunswick [etc.]: Rutgers University Press.
- Allemann-Ghionda, C. (2004). *Einführung in die vergleichende Erziehungswissenschaft* (Studium Paedagogik). Weinheim: Beltz.
- Anbuhl, M. (2016). *Kein Anschluss mit diesem Abschluss? DGB-Expertise zu den Chancen von Jugendlichen mit Hauptschulabschluss auf dem Ausbildungsmarkt*. Eine Analyse anhand der Zahlen der DIHK-Lehrstellenbörse vom 13. Juli 2016 und des Berichts „Bildung in Deutschland 2016“ (DBG-Bundesvorstand, Hrsg.), Berlin. <https://www.google.com/url?sa=t&rct=j&q=&esrc=KHbs8A0YQFggxMAE&url=https://www.dgb.de/2Fthemen/2F%2B%2Bco%2B%2Baff65d8e-5247-11e6-9b99-525400e5a74a&usg=AOvVaw0ydY9rBe208iQfzm17lWFk>. Zugegriffen: 5. März 2018.
- Anweiler, O. (1969). Probleme eines Leistungsvergleichs der Bildungssysteme in der Bundesrepublik und der DDR. *Bildung und Erziehung* 22, 241–253. doi:10.7788/bue-1969-jg25
- Artelt, C., Brunner, M., Schneider, W., Prenzel, M. & Neubrand, M. (2003). Literacy oder Lehrplanvalidität? Ländervergleiche auf der Basis lehrplanoptimierter PISA-Tests. In J. Baumert, C. Artelt, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, K.-J. Tillmann & M. Weiß (Hrsg.), *PISA 2000 — Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland* (S. 77–108). Wiesbaden: VS Verlag für Sozialwissenschaften. doi:10.1007/978-3-322-97590-4_4
- Austan-Smith, D. & Fyer, R. G. (2005). An economic analysis of ‘acting white’. *Quarterly Journal of Economics* 120, 551–583. doi:10.1093/qje/120.2.551
- Autorengruppe Bildungsberichterstattung. (2016). *Bildung in Deutschland 2016. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration*. Bielefeld: wbv.
- Autorengruppe Bildungsberichterstattung. (2018). *Bildung in Deutschland 2018. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration*. Bielefeld: wbv.
- AYala, C. C., Shavelson, R. J., Araceli Ruiz-Primo, M., Brandon, P. R., Yin, Y., Furtak, E. M., Young, D. B. & Tomita, M. K. (2008). From Formal Embedded Assessments

- to Reflective Lessons. The Development of Formative Assessment Studies. *Applied Measurement in Education* 21 (4), 315–334. doi:10.1080/08957340802347787
- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2016). *Multivariate Analysemethoden* (14. Aufl. 2016). Berlin, Heidelberg: Springer. doi:10.1007/978-3-662-46076-4
- Baron, R. M. & Kenny, D. A. (1986). The moderator-mediator Variable distinction in social psychology research. Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology* 51 (6), 1173–1182. doi:10.1037/0022-3514.51.6.1173
- Barrett, P. (1986). Factor comparisons. An examination of three methods. *Personality and Individual Differences* 7 (3), 327–340. doi:10.1016/0191-8869(86)90008-5
- Bastian, J., Helsper, W., Reh, S. & Schelle, C. (2000). *Professionalisierung im Lehrberuf. Von der Kritik der Lehrerrolle zur pädagogischen Professionalität* (Studien zur Schul- und Bildungsforschung, Bd. 12). Opladen: Leske + Budrich.
- Baumert, J., Cortina, K. S. & Leschinsky, A. (2008). Grundlegende Entwicklungen und Strukturprobleme im allgemeinbildenden Schulwesen. In K. S. Cortina, J. Baumert, A. Leschinsky, K. U. Mayer & L. Trommer (Hrsg.), *Das Bildungswesen in der Bundesrepublik Deutschland* (S. 53–130). Reinbek bei Hamburg: Rowohlt-Taschenbuch-Verlag.
- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft* 9 (4), 469–520. doi:10.1007/s11618-006-0165-2
- Baumert, J. & Kunter, M. (2011). Das Kompetenzmodell von COACTIV. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 29–53). Münster: Waxmann.
- Baumert, J., Kunter, M., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (2011). Professionelle Kompetenz von Lehrkräften, kognitiv aktivierender Unterricht und die mathematische Kompetenz von Schülerinnen und Schülern (COACTIV). Ein Forschungsprogramm. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 7–25). Münster: Waxmann.
- Baumert, J., Kunter, M., Brunner, M., Krauss, S., Blum, W. & Neubrand, M. (2004). Mathematikunterricht aus Sicht der PISA-Schülerinnen und Schüler und ihrer Lehrkräfte. In PISA-Konsortium (Hrsg.), *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland. Ergebnisse des zweiten internationalen Vergleichs* (S. 314–354). Münster: Waxmann.
- Baumert, J., Lehmann, R., Lehrke, M., Schmitz, B., Clausen, M., Hosenfeld, I. & Köller, O. (1997). *TIMSS. Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich*. Opladen: Leske und Budrich. doi:10.1007/978-3-322-95096-3
- Baumert, J., Lehmann, R., Lehrke, M., Schmitz, B., Clausen, M., Hosenfeld, I., Köller, O. & Neubrand, M. (1997). *TIMSS. Mathematisch-Naturwissenschaftlicher Unterricht im internationalen Vergleich*. Zusammenfassung deskriptiver Ergebnisse, Berlin. doi:10.1007/978-3-322-95096-3
- Baumert, J., Maaz, K., Neumann, M., Becker, M. & Dumont, H. (2017). Die Berliner Strukturreform. Hintergründe, Zielstellungen und theoretischer Rahmen. In M. Neumann, M. Becker, J. Baumert, K. Maaz & O. Köller (Hrsg.), *Zweigliedrigkeit im*

- deutschen Schulsystem. *Potenziale und Herausforderungen in Berlin* (S. 9–39). Münster: Waxmann.
- Baumert, J., Stanat, P. & Watermann, R. (Hrsg.). (2006a). *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit*. Wiesbaden: VS Verlag für Sozialwissenschaften. doi:10.1007/978-3-531-90082-7
- Baumert, J., Stanat, P. & Watermann, R. (2006b). Schulstruktur und die Entstehung differenzieller Lern- und Entwicklungsmilieus. In J. Baumert, P. Stanat & R. Watermann (Hrsg.), *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit* (S. 95–188). Wiesbaden: VS Verlag für Sozialwissenschaften. doi:10.1007/978-3-531-90082-7_4
- Baumert, J., Trautwein, U. & Artelt, C. (2003). Schulumwelten. Institutionelle Bedingungen des Lehrens und Lernens. In J. Baumert, C. Artelt, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, K.-J. Tillmann & M. Weiß (Hrsg.), *PISA 2000 — Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland* (S. 261–331). Wiesbaden: VS Verlag für Sozialwissenschaften. doi:10.1007/978-3-322-97590-4_11
- Baumgartner, H. & Steenkamp, J.-B. E. M. (2001). Response styles in marketing research. A cross-national investigation. *Journal of Marketing Research* 38 (2), 143–156. doi:10.1509/jmkr.38.2.143.18840
- Bayer, S., Vieluf, S. & Klieme, E. (2015). *Unterrichtsqualität. Schulformen im Vergleich*. Präsentation auf der 80. Tagung Arbeitsgruppe für empirische pädagogische Bildungsforschung (AEPF) in Göttingen.
- Bayerisches Staatsministerium für Unterricht und Kultus. (2007). *Gymnasialschulordnung. GSO*. <http://www.gesetze-bayern.de/Content/Document/BayGSO>.
- Becker, M., Lüdtke, O., Trautwein, U. & Baumert, J. (2006). Leistungszuwachs in Mathematik. Evidenz für einen Schereneffekt im mehrgliedrigen Schulsystem? *Zeitschrift für Pädagogische Psychologie* 20 (4), 233–242. doi:10.1024/1010-0652.20.4.233
- Becker, M., Lüdtke, O., Trautwein, U., Köller, O. & Baumert, J. (2012). The differential effects of school tracking on psychometric intelligence. Do academic-track schools make students smarter? *Journal of Educational Psychology* 104 (3), 682–699. doi:10.1037/a0027608
- Becker, M. & Neumann, M. (2016). Context-related changes in academic self concept development. On the long-term persistence of big-fish-little-pond effects. *Learning and Instruction* 45, 31–39. doi:10.1016/j.learninstruc.2016.06.003
- Becker, M. & Neumann, M. (2018). Longitudinal big-fish-little-pond effects on academic self-concept development during the transition from elementary to secondary schooling. *Journal of Educational Psychology* 110 (6), 882–897. doi:10.1037/edu0000233
- Bellenberg, G. & Klemm, K. (2011). Die Grundschule im deutschen Schulsystem. In W. Einsiedler (Hrsg.), *Handbuch Grundschulpädagogik und Grundschuldidaktik* (UTB, Bd. 8444, 3., vollst. überarb. Aufl., S. 45–51). Bad Heilbrunn: Klinkhardt.
- Bellin, N. (2009). *Klassenkomposition, Migrationshintergrund und Leistung. Mehrebenenanalysen zum Sprach- und Leseverständnis von Grundschulern*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Berliner, D. C. (2016). The near impossibility of testing for teacher quality. *Journal of Teacher Education* 56 (3), 205–213. doi:10.1177/0022487105275904

- Beuckelaer, A. de (2005). *Measurement invariance issues in international management research*. Limburgs Universitair Centrum, Limburg, Niederlande.
- Bierhoff, H.-W. (2002). *Einführung in die Sozialpsychologie* (Beltz Studium). Weinheim [u.a.]: Beltz.
- Bildungsserver Mecklenburg-Vorpommern. (2018). *Regionale Schule*, Bildungsserver Mecklenburg-Vorpommern. <https://www.bildung-mv.de/schueler/schule-und-unterricht/schularten/regionale-schule/>. Zugegriffen: 27. März 2018.
- Black, P. & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice* 5 (1), 7–74. doi:10.1080/0969595980050102
- Black, P. & Wiliam, D. (1998b). Inside the black box. Raising standards through classroom assessment. *Phi Delta Kappa* 80 (2), 139–147.
- Blömeke, S. (2009a). Lehrerbildung. In S. Blömeke, T. Bohl, L. Haag, G. Lang-Wojtask & W. Sacher (Hrsg.), *Handbuch Schule. Theorie – Organisation – Entwicklung* (UTB, Bd. 8392, S. 483–490). Bad Heilbrunn: Klinkhardt.
- Blömeke, S. (2009b). Lehrerbildung in Deutschland. *PÄD-Forum: unterrichten erziehen* 37/38 (1), 5–8.
- Blömeke, S. (2019). Lehrerbildung. In O. Köller, M. Hasselhorn, F. W. Hesse, K. Maaz, J. Schrader, H. Solga, C. K. Spieß & K. Zimmer (Hrsg.), *Das Bildungswesen in Deutschland. Bestand und Potenziale* (1. Aufl.). Klinkhardt.
- Blömeke, S. & Müller, C. (2009). Zum Zusammenhang von Allgemeiner Didaktik und Lehr-Lernforschung im Unterrichtsgeschehen. In M. A. Meyer, S. Hellekamps & M. Prenzel (Hrsg.), *Perspektiven der Didaktik* (Zeitschrift für Erziehungswissenschaft Sonderheft, Bd. 9, S. 239–258). Wiesbaden: VS Verlag für Sozialwissenschaften. doi:10.1007/978-3-531-91775-7_16
- Blömeke, S., Kaiser, G. & Lehmann, R. (Hrsg.). (2008). *Professionelle Kompetenz angehender Lehrerinnen und Lehrer. Wissen, Überzeugungen und Lerngelegenheiten deutscher Mathematikstudierender und -referendare. Erste Ergebnisse zur Wirksamkeit der Lehrerbildung*. Münster: Waxmann.
- Blömeke, S., König, J., Busse, A., Suhl, U., Benthien, J., Döhrmann, M. & Kaiser, G. (2014). Von der Lehrerbildung in den Beruf. Fachbezogenes Wissen als Voraussetzung für Wahrnehmung, Interpretation und Handeln im Unterricht. *Zeitschrift für Erziehungswissenschaft* 17 (3), 509–542. doi:10.1007/s11618-014-0564-8
- Blum, W., Neubrand, M., Ehmke, T., Senkbeil, M., Jordan, A., Ulfig, F. & Carstensen, C. H. (2004). Mathematische Kompetenz. In PISA-Konsortium (Hrsg.), *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland. Ergebnisse des zweiten internationalen Vergleichs* (S. 47–92). Münster: Waxmann.
- Bogner, K. & Landrock, U. (2015). *Antworttendenzen in standardisierten Umfragen* (GESIS Survey Guidelines). Mannheim: GESIS – Leibniz Institut für Sozialwissenschaften.
- Bohl, T. (2000). *Unterrichtsmethoden in der Realschule. Eine empirische Untersuchung zum Gebrauch ausgewählter Unterrichtsmethoden an staatlichen Realschulen in Baden-Württemberg; ein Beitrag zur deskriptiven Unterrichtsmethodenforschung* (1. Aufl.). Bad Heilbrunn/Obb.: Klinkhardt (Zugl.: Tübingen, Univ., Diss., 2000).
- Bohl, T., Kleinknecht, M., Batzel, A. & Richery, P. (2013). *Aufgabenkultur in der Schule. Eine vergleichende Analyse von Aufgaben und Lehrerhandeln im Hauptschul-, Realschul- und Gymnasialunterricht* (Schul- und Unterrichtsforschung, Bd. 15, 2. unveränd. Aufl.). Baltmannsweiler: Schneider-Verl. Hohengehren.

- Böhme, J. (2000). *Schulmythen und ihre imaginäre Verbürgung durch oppositionelle Schüler. Ein Beitrag zur Etablierung erziehungswissenschaftlicher Mythosforschung* (1. Aufl.). Bad Heilbrunn/Obb.: Klinkhardt.
- Bohnsack, R. (2007). *Rekonstruktive Sozialforschung. Einführung in qualitative Methoden* (UTB Erziehungswissenschaft, Sozialwissenschaft, Bd. 8242, 8., durchges. Aufl.). Opladen: Budrich.
- Bortz, J. & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler* (Springer-Lehrbuch, 7. vollst. überarb. und erw. Aufl.). Berlin [u.a.]: Springer. doi:10.1007/978-3-642-12770-0
- Böttcher, W. (2003). Kerncurricula und die Steuerung der allgemeinbildenden Schulen. In H.-P. Füssel & P. M. Roeder (Hrsg.), *Recht – Erziehung – Staat. Zur Genese einer Problemkonstellation und zur Programmatik ihrer zukünftigen Entwicklung* (Zeitschrift für Pädagogik, Beiheft; 47, S. 215–233). Weinheim: Beltz.
- Bourdieu, P. (1983). Ökonomisches Kapital, kulturelles Kapital, soziales Kapital. In R. Kreckel (Hrsg.), *Soziale Ungleichheiten* (Soziale Welt. Sonderband, Bd. 2, S. 183–199). Göttingen: Schwartz.
- Bourdieu, P. (2001). *Wie die Kultur zum Bauern kommt. Über Bildung, Schule und Politik* (Schriften zur Politik & Kultur, Bd. 4). Hamburg: VSA-Verlag.
- Bourdieu, P. & Passeron, J.-C. (1971). *Die Illusion der Chancengleichheit. Untersuchungen zur Soziologie des Bildungswesens am Beispiel Frankreichs* (Texte und Dokumente zur Bildungsforschung, 1. Aufl.). Stuttgart: Klett.
- Breidenstein, G. (2010). Überlegungen zu einer Theorie des Unterrichts. *Zeitschrift für Pädagogik* 56 (6), 869–871.
- Breidenstein, G. & Kelle, H. (2002). Die Schulklasse als Publikum. Zum Verhältnis von Peer Culture und Unterricht. *Die Deutsche Schule* 94 (3), 318–329.
- Briggs, S. R. & Cheek, J. M. (1986). The role of factor analysis in the development and evaluation of personality scales. *Journal of Personality and Social Psychology* 54 (1), 106–148. doi:10.1111/j.1467-6494.1986.tb00391.x
- Bromme, R. (1997). Kompetenzen, Funktionen und unterrichtliches Handeln des Lehrers. In F. E. Weinert (Hrsg.), *Enzyklopädie der Psychologie. Pädagogische Psychologie*. Psychologie des Lernens und der Schule (Bd. 3, S. 177–212). Göttingen: Hogrefe.
- Bromme, R., Rheinberg, F., Minsel, B., Winteler, A. & Weidenmann, B. (2006). Die Erziehenden und Lehrenden. In A. Krapp & B. Weidenmann (Hrsg.), *Pädagogische Psychologie* (S. 296–355). Weinheim: Beltz.
- Brookhart, S. M., Moss, C. M. & Long, B. A. (2008). Formative assessment that empowers. *Educational Leadership* 66 (3), 52–57.
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research* (Methodology in the social sciences). New York: Guilford Press.
- Bundesinstitut für Berufsbildung (Hrsg.). (2016). *Datenreport zum Berufsbildungsbericht 2015. Informationen und Analysen zur Entwicklung der beruflichen Bildung*. Bonn: BIBB Bundesministerium für Berufsbildung.
- Bundschuh, K. & Baier, S. (2009). Sonderschulwesen und sonderpädagogische Einrichtungen. In S. Blömeke, T. Bohl, L. Haag, G. Lang-Wojtask & W. Sacher (Hrsg.), *Handbuch Schule. Theorie – Organisation – Entwicklung* (UTB, Bd. 8392, S. 288–298). Bad Heilbrunn: Klinkhardt.
- Campbell, D. T., Kenny, D. A. & Reichardt, C. S. (1999). *A primer on regression artifacts* (Methodology in the social sciences). New York: Guilford Press.

- Cannell, C., Miller, P. & Oksenberg, L. (1981). Research on interviewing techniques. In S. Leinhardt (Hrsg.), *Sociological methodology* (Social and behavioral science series, S. 389–437). San Francisco: Jossey-Bass. doi:10.2307/270748
- Carlsen, W. S. (1987). *Why do you ask? The effect of science teacher subject-matter knowledge on teacher questioning and classroom discourse*. Paper presented at the annual meeting of the American Educational Research Association (April 1987), Washington, DC.
- Carroll, J. B. (1963). A model of school learning. *Teachers College Record* 64 (8), 723–733.
- Carroll, J. B. (1989). The Carroll model. A 25-year retrospective and prospective view. *Educational Researcher* 18 (1), 26–31. doi:10.3102/0013189X018001026
- Carvalho, I. N. (2017). *Gymnasium und Ganztagschule. Videographische Fallstudie zur Konstitution pädagogischer Ordnung*. Wiesbaden: Springer Fachmedien Wiesbaden. doi:10.1007/978-3-658-18358-5
- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research* 1, 245–276. doi:10.1207/s15327906mbr0102_10
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling* 14 (3), 464–504. doi:10.1080/10705510701301834
- Cheung, G. W. & Rensvold, R. B. (1999). Testing factorial invariance across groups. A reconceptualization and proposed new method. *Journal of Management* 25 (1), 1–27. doi:10.1177/014920639902500101
- Cheung, G. W. & Rensvold, R. B. (2000). Assessing extreme and acquiescence response sets in cross-cultural research using structural equations modeling. *Journal of Cross-Cultural Psychology* 31 (2), 187–212. doi:10.1177/0022022100031002003
- Clausen, M. (2002). *Unterrichtsqualität. Eine Frage der Perspektive* (Pädagogische Psychologie und Entwicklungspsychologie, Bd. 29). Münster [u.a.]: Waxmann.
- Cohen, J., McCabe, E. M., Michelli, N. M. & Pickeral, T. (2009). School Climate. Research, policy, practice and teacher education. *Teachers College Record* 111 (1), 180–213.
- Comenius, J. A. (2000). *Große Didaktik* (9., Aufl.) (Andreas Flitner., Hrsg.) (Pädagogische Texte). Stuttgart: Klett-Cotta, Original 1657.
- Costello, A. B. & Osborne, J. W. (2005). Best practice in Exploratory Factor Analysis. Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation* 10 (7). <http://pareonline.net/getvn.asp?v=10&n=7>. Zugriffen: 25. Oktober 2017.
- Crosnoe, R. (2009). Low-income students and the socioeconomic composition of public high schools. *American sociological review* 74 (5), 709–730. doi:10.1177/000312240907400502
- Davidov, E. (2008). A cross-country and cross-time comparison of the Human Values Measurement with the second round of the European Social Survey. *Survey Research Methods* (2008) 2 (1), 33–46.
- Deci, E. L. & Ryan, R. M. (1996). *Intrinsic motivation and self-determination in human behavior* (Perspectives in social psychology, [5]th print). New York: Plenum Press.
- Decristan, J., Klieme, E., Kunter, M., Hochweber, J., Büttner, G., Fauth, B., Hondrich, L. A., Rieser, S., Hertelt, S., Hardy, I., Büttner, G., Hondrich, A. L., Hertel, S. & Hardy, I. (2015). Embedded formative assessment and classroom process quality.

- How do they interact in promoting science understanding? *American Educational Research Journal* 52 (6), 1133–1159. doi:10.3102/0002831215596412
- Derting, T. L. & Ebert-May, D. (2010). Learner-centered inquiry in undergraduate biology. Positive relationships with long-term student achievement. *CBE life sciences education* 9 (4), 462–472. doi:10.1187/cbe.10-02-0011
- DESI-Konsortium (Hrsg.). (2008). *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Zentrale Befunde der Studie Deutsch Englisch Schülerleistung International (DESI)*. Weinheim [u.a.]: Beltz.
- Diederich, J. & Tenorth, H.-E. (1997). *Theorie der Schule. Ein Studienbuch zu Geschichte, Funktionen und Gestaltung*. Berlin: Cornelsen Scriptor.
- Diehl, J. M. (1983). *Varianzanalyse* (Methoden in der Psychologie, Bd. 3, 4., unveränd. Aufl.). Frankfurt am Main: Fachbuchhandlung für Psychologie, Verl.-Abt.
- Dietrich, J., Dicke, A.-L., Kracke, B. & Noack, P. (2015). Teacher support and its influence on students' intrinsic value and effort. Dimensional comparison effects across subjects. *Learning and Instruction* 39, 45–54. doi:10.1016/j.learninstruc.2015.05.007
- DIHK. (2017). *Ausbildung 2017. Ergebnisse einer DIHK-Online-Unternehmensbefragung*, Berlin.
- DIPF; Universität Bielefeld; Universität Bremen; IQB. (2018). *Bericht der Expertengruppe der Evaluation der Bremer Schulreform* (Expertengruppe zur Evaluation der Bremer Schulreform, Hrsg.). <https://www.dipf.de/de/forschung/aktuelle-projekte/pdf/steubis/bremen-evaluation-maerz-2018-gesamtbericht>. Zugegriffen: 12. April 2018.
- Ditton, H. (2000). Qualitätskontrolle und Qualitätssicherung in Schule und Unterricht. Ein Überblick zum Stand der empirischen Forschung. In A. Helmke, W. Hornstein & E. Terhart (Hrsg.), *Qualität und Qualitätssicherung im Bildungsbereich. Schule, Sozialpädagogik, Hochschule*. Zeitschrift für Pädagogik, Beiheft 41) (S. 73–92). Weinheim: Beltz.
- Ditton, H. & Kreyer, L. (1995). Qualität von Schule und Unterricht. Empirische Befunde zur Fragestellung und Aufgaben der Forschung. *Zeitschrift für Pädagogik* 41 (4), 507–529.
- Ditton, H. & Krüken, J. (2010). Bildungslaufbahnen im differenzierten Schulsystem. Entwicklungsverläufe von Laufbahneempfehlungen und Bildungsaspirationen in der Grundschulzeit. In J. Baumert, K. Maaz & U. Trautwein (Hrsg.), *Bildungsentscheidungen. Zeitschrift für Erziehungswissenschaft Sonderheft 12/2009* (S. 74–102). Wiesbaden: VS Verlag für Sozialwissenschaften. doi:10.1007/978-3-531-92216-4_4
- Dohrmann, J., Feldhoff, T., Steinert, B. & Klieme, E. (2019). Überzeugungen von Lehrkräften, Adaptivität des Unterrichts und Lernergebnisse im Fach Englisch. *Zeitschrift für Pädagogik* 65 (1), 56–72.
- Dolch, J. (1982). *Lehrplan des Abendlandes. Zweieinhalb Jahrtausende seiner Geschichte* (Unveränd. reprograf. Nachdr. d. 3. Aufl.). Darmstadt: Wissenschaftliche Buchgesellschaft.
- Drechsel, B. & Senkbeil, M. (2004). Institutionelle Rahmenbedingungen von Schule und Unterricht. In PISA-Konsortium (Hrsg.), *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland. Ergebnisse des zweiten internationalen Vergleichs* (S. 284–291). Münster: Waxmann.
- Dreeben, R. & Barr, R. (1988). Classroom composition and the design of instruction. *Sociology of Education* 61 (3), 129–142. doi:10.2307/2112622

- Drewek, P. (2015). Das gegliederte Schulwesen in Deutschland im historischen Prozess. Ansätze, Quellen und Desiderate der historischen Bildungsforschung. *Archivpflege in Westfalen-Lippe* 83, 5–10.
- Drewek, P. & Tenorth, H.-E. (2001). Das deutsche Bildungswesen im 19. und 20. Jahrhundert. Systemdynamik und Systemreflektion. In H. J. Apel (Hrsg.), *Das öffentliche Bildungswesen. Historische Entwicklung, gesellschaftliche Funktionen, pädagogischer Streit* (S. 49–79). Bad Heilbrunn/Obb.: Klinkhardt.
- Dubberke, T., Kunter, M., McElvany, N., Brunner, M. & Baumert, J. (2008). Lerntheoretische Überzeugungen von Mathematiklehrkräften. *Zeitschrift für Pädagogische Psychologie* 22 (34), 193–206. doi:10.1024/1010-0652.22.34.193
- Dühlmeier, B. (2009). Schul- und Bildungsreformen im deutschsprachigen Raum seit 1945. In S. Blömeke, T. Bohl, L. Haag, G. Lang-Wojtask & W. Sacher (Hrsg.), *Handbuch Schule. Theorie – Organisation – Entwicklung* (UTB, Bd. 8392, S. 162–170). Bad Heilbrunn: Klinkhardt.
- Dumont, H., Neumann, M., Maaz, K. & Trautwein, U. (2013). Die Zusammensetzung der Schülerschaft als Einflussfaktor für Schulleistungen. *Psychologie in Erziehung und Unterricht* 60 (3), 163–183. doi:10.2378/peu2013.art14d
- Durkheim, E. (1962). *The rules of sociological method*. New York: Free Press.
- Eder, D. & Sanford, S. (1986). The development and maintenance of interactional norms among early adolescents. In P. A. Adler & P. Adler (Hrsg.), *Sociological studies of child development. A research annual* (Sociological studies of child development, v. 1 (1986)). Greenwich, Conn.: Jai Press.
- Eder, F. (2010). Schul- und Klassenklima. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (4., überarb. und erw. Aufl., S. 694–703). Weinheim: Beltz.
- Eder, F. & Hörl, G. (Hrsg.). (2010). *Schule auf dem Prüfstand. Hauptschule und gymnasiale Unterstufe im Spiegel der Forschung* (Austria: Forschung und Wissenschaft / Erziehungswissenschaft, Bd. 10). Wien: Lit.
- Ehmke, T., Köller, O. & Stanat, P. (2017). Äquivalenz der Erfassung mathematischer Kompetenzen in PISA 2012 und im IQB-Ländervergleich 2012. *Zeitschrift für Erziehungswissenschaft* 20 (S2), 37–59. doi:10.1007/s11618-017-0751-5
- Eickhorst, A. (2011). Das Unterrichtsverständnis der empirischen Lehr-Lern-Forschung. In W. Meseth, M. Proske & F.-O. Radtke (Hrsg.), *Unterrichtstheorien in Forschung und Lehre* (S. 50–66). Bad Heilbrunn: Verlag Julius Klinkhardt.
- Einsiedler, W. (1997). Unterrichtsqualität und Leistungsentwicklung. Literaturüberblick. In Weinert, F. E. Weinert & A. Helmke (Hrsg.), *Entwicklung im Grundschulalter* (225-140). Weinheim: Beltz Psychologie-Verl.-Union.
- Elen, J., Clarebout, G., Léonard, R. & Lowyck, J. (2007). Student-centred and teacher-centred learning environments. What students think. *Teaching in Higher Education* 12 (1), 105–117. doi:10.1080/13562510601102339
- Esser, H. (2006). *Sprache und Integration. Die sozialen und Bedingungen und Folgen des Spracherwerbs von Migranten*. Frankfurt am Main: Campus.
- Esser, H. (2009). Der Streit um die Zweisprachigkeit. Was bringt die Bilingualität? In I. Gogolin & U. Neumann (Hrsg.), *Streitfall Zweisprachigkeit – the bilingualism controversy* (S. 69–88). Wiesbaden: VS Verlag für Sozialwissenschaften. doi:10.1007/978-3-531-91596-8_5
- Fahrmeir, A. (2017). *Deutsche Geschichte* (Beck'sche Reihe, Bd. 2875, 1. Aufl.). München: C.H. Beck. doi:10.17104/9783406715143

- Fauth, B., Decristan, J., Rieser, S., Klieme, E. & Buttner, G. (2014). Student ratings of teaching quality in primary school. Dimensions and prediction of student outcomes. *Learning and Instruction* 29, 1–9. doi:10.1016/j.learninstruc.2013.07.001
- Fees, K. (2009). Realschule. In S. Blömeke, T. Bohl, L. Haag, G. Lang-Wojtask & W. Sacher (Hrsg.), *Handbuch Schule. Theorie – Organisation – Entwicklung* (UTB, Bd. 8392, S. 245–253). Bad Heilbrunn: Klinkhardt.
- Fend, H. (1982). *Gesamtschule im Vergleich. Bilanz der Ergebnisse des Gesamtschulversuchs*. Weinheim: Beltz.
- Fend, H. (1988). Schulqualität. Die Wiederentdeckung der Schule als Gestaltungsebene. *Neue Sammlung* 28 (4).
- Fend, H. (1989). „Pädagogische Programme“ und ihre Wirksamkeit. Das Beispiel der Umdeutung schulischer Normen und Erwartungen in der Altersgruppe. In W. Breyvogel (Hrsg.), *Pädagogische Jugendforschung. Erkenntnisse und Perspektiven* (S. 187–209). Wiesbaden: VS Verlag für Sozialwissenschaften. doi:10.1007/978-3-322-97188-3_12
- Fend, H. (2008). *Schule gestalten. Systemsteuerung, Schulentwicklung und Unterrichtsqualität*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Fend, H. (2009). *Neue Theorie der Schule. Einführung in das Verstehen von Bildungssystemen* (2., durchgesehene Auflage). Wiesbaden: VS Verlag für Sozialwissenschaften. doi:10.1007/978-3-531-91788-7
- Fend, H. & Specht, W. (1977). Zur Lage der leistungsschwächeren Schüler in unterschiedlichen Schulsystemen. In Redaktion Betrifft: Erziehung (Hrsg.), *Hauptschule* (S. 27–55). Weinheim: Beltz.
- Festinger, L. (1950). Informal social communication. *Psychological Review* 57 (5), 272–282. doi:10.1037/h0056932
- Fives, H. & Buehl, M. M. (2012). Spring cleaning for the „messy“ construct of teachers’ beliefs: What are they? Which have been examined? What can they tell us? In K. R. Harris, S. Graham, T. Urdan, S. Graham, J. M. Royer & M. Zeidner (Hrsg.), *APA educational psychology handbook, Vol 2: Individual differences and cultural and contextual factors* (S. 471–499). Washington: American Psychological Association. doi:10.1037/13274-019
- Fredrick, W. C. & Walberg, H. J. (2014). Learning as a function of time. *The Journal of Educational Research* 73 (4), 183–194. doi:10.1080/00220671.1980.10885233
- Fromm, S. (2010). *Datenanalyse mit SPSS für Fortgeschrittene 2. Multivariate Verfahren für Querschnittsdaten* (Springer-Lehrbuch). Wiesbaden: VS Verlag für Sozialwissenschaften. doi:10.1007/978-3-531-92026-9
- Gailberger, S. & Willenberg, H. (2008). Leseverstehen Deutsch. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Zentrale Befunde der Studie Deutsch Englisch Schülerleistung International (DESI)* (S. 60–71). Weinheim [u.a.]: Beltz.
- Gamoran, A., Nystrand, M., Berends, M. & LePore, L. (1995). An organizational analysis of the effects of ability grouping. *American Educational Research Journal* 32 (4), 687–715. doi:10.3102/00028312032004687
- Gaude, P. (1996). Hat die Gesamtschule das traditionelle Schulsystem verändert? In H. Gudjons & A. Köpke (Hrsg.), *25 Jahre Gesamtschule in der Bundesrepublik Deutschland. Eine bildungspolitische und pädagogische Bilanz* (S. 84–100). Bad Heilbrunn: Klinkhardt.

- Gegern, K. J. (1995). Social construction and the educational process. In L. P. Steffe & J. Gale (Hrsg.), *Constructivism in education* (S. 17–39). NJ: Taylor and Francis.
- Gehrmann, A. (2016). „Die Systemfrage kann als relativ abschließend behandelbar angesehen werden“. Anmerkungen zu Schulentwicklung, Bildungsexpansion und Lehrerbedarf nach 1945. In T.-S. Idel, F. Dietrich, K. Kunze, K. Rabenstein & A. Schütz (Hrsg.), *Professionsentwicklung und Schulstrukturereform. Zwischen Gymnasium und neuen Schulformen in der Sekundarstufe* (Studien zur Professionsforschung und Lehrerbildung, S. 23–46). Bad Heilbrunn: Klinkhardt.
- Gellert, U. (1998). *Von Lernerfahrungen zu Unterrichtskonzeptionen. Eine soziokulturelle Analyse von Vorstellungen angehender Lehrerinnen und Lehrer zu Mathematik und Mathematikunterricht* (Akademische Abhandlungen zu den Erziehungswissenschaften, 1. Aufl.). Berlin: VWF, Verl. für Wiss. und Forschung. doi:10.1007/BF03338872
- Gemeinnützige Gesellschaft Gesamtschule. (2015). *Grundsatzposition* (Gemeinnützige Gesellschaft Gesamtschule, Hrsg.). <http://ggg-bund.de/index.php/service/downloads/category/1-ggg-positionen?download=409:ggg-grundsatzposition-2015>. Zugegriffen: 22. März 2018.
- Geulen, D. (2007). Sozialisation. In H. Joas (Hrsg.), *Lehrbuch der Soziologie* (3., überarb. und erw. Aufl., S. 137–158). Frankfurt/Main [u.a.]: Campus-Verl.
- Giesecke, H. (2002). Fächer, Stoffe, Bildung. In W. Böttcher & P. E. Kalb (Hrsg.), *Kerncurriculum. Was Kinder in der Grundschule lernen sollen; eine Streitschrift* (Beltz Pädagogik, S. 64–81). Weinheim: Beltz.
- Giesinger, J. (2016). Bildung im liberalen Staat. Von Humboldt zu Rawls. In R. Casale, H.-C. Koller & N. Ricken (Hrsg.), *Das Pädagogische und das Politische. Zu einem Topos der Erziehungs- und Bildungsphilosophie* (Schriftenreihe der Kommission Bildungs- und Erziehungsphilosophie in der Deutschen Gesellschaft für Erziehungswissenschaft, S. 27–40). Paderborn: Ferdinand Schöningh. doi:10.30965/9783657782680_004
- Grimm, S. D. & Church, T. A. (1999). A cross-cultural study of response bias in personality measures. *Journal of Research in Personality* 33, 415–441. doi:10.1006/jrpe.1999.2256
- Gruehn, S. (2000). *Unterricht und schulisches Lernen. Schüler als Quellen der Unterrichtsbeschreibung* (Pädagogische Psychologie und Entwicklungspsychologie, Bd. 12). Münster: Waxmann.
- Gruschka, A. (2013). *Unterrichten. Eine pädagogische Theorie auf empirischer Basis*. Opladen: Verlag Barbara Budrich. doi:10.2307/j.ctvdf0d0p
- Guill, K., Gröhlich, C., Scharenberg, K., Wendt, H. & Bos, W. (2010). Die mathematische Kompetenz der Schülerinnen und Schüler. In W. Bos (Hrsg.), *KESS 8. Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 8* (HANSE – Hamburger Schriften zur Qualität im Bildungswesen, Bd. 6, S. 37–48). Münster: Waxmann.
- Guill, K., Lüdtke, O. & Köller, O. (2017). Academic tracking is related to gains in students' intelligence over four years. Evidence from a propensity score matching study. *Learning and Instruction* 47, 43–52. doi:10.1016/j.learninstruc.2016.10.001
- Guttman, L. (1954). Some necessary conditions for common factor analysis. *Psychometrika* 19, 149–161. doi:10.1007/BF02289162

- Haag, L. & Hopperdietzel, H. (2009). Gymnasium. In S. Blömeke, T. Bohl, L. Haag, G. Lang-Wojtask & W. Sacher (Hrsg.), *Handbuch Schule. Theorie – Organisation – Entwicklung* (UTB, Bd. 8392, S. 253–261). Bad Heilbrunn: Klinkhardt.
- Hage, K., Bischoff, H., Dichanz, H., Eubel, K.-D., Oehlschläger, H.-J. & Schwittmann, D. (1985). *Das Methoden-Repertoire von Lehrern. Eine Untersuchung zum Schulalltag der Sekundarstufe I*. Opladen: Leske und Budrich. doi:10.1007/978-3-663-09543-9
- Hameyer, U., Frey, K., Haft, H. & Hameyer (Hrsg.). (1983). *Handbuch der Curriculumforschung. Erste Ausgabe – Übersichten zur Forschung 1970–1981*. Weinheim: Beltz.
- Handl, A. (2010). *Multivariate Analysemethoden. Theorie und Praxis multivariater Verfahren unter besonderer Berücksichtigung von S-PLUS* (Statistik und ihre Anwendungen, 2. Aufl.). Heidelberg: Springer.
- Hänsel, D. (2005). Die Historiographie der Sonderschule. Eine kritische Analyse. *Zeitschrift für Pädagogik* 51 (1), 101–115.
- Harker, R. & Tymms, P. (2004). The effects of student composition on school outcomes. *School Effectiveness and School Improvement* 15 (2), 177–199. doi:10.1076/sesi.15.2.177.30432
- Harris, D. N. (2010). How do school peers influence student educational outcomes? Theory and evidence from economics and other social sciences. *Teachers College Record* 112 (4), 1163–1197.
- Harsch, C., Schröder, K. & Neumann, A. (2008). Schreiben Englisch. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Zentrale Befunde der Studie Deutsch Englisch Schülerleistung International (DESI)* (S. 139–148). Weinheim [u.a.]: Beltz.
- Hashweh, M. Z. (1987). Effects of subject-matter knowledge in the teaching of biology and physics. *Teaching and Teacher Education* 3 (2), 109–120. doi:10.1016/0742-051X(87)90012-6
- Hattie, J. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement* (Educational research, 1. publ). doi:10.4324/9780203887332
- He, J. & van de Vijver, F. (2012). Bias and Equivalence in Cross-Cultural Research. *Online Readings in Psychology and Culture* 2 (2). doi:10.9707/2307-0919.1111
- Heine, J.-H., Nagy, G., Meinck, S., Zühlke, O. & Mang, J. (2017). Empirische Grundlage, Stichprobenausfall und Adjustierung im PISA-Längsschnitt 2012–2013. *Zeitschrift für Erziehungswissenschaft* 20 (S2), 287–306. doi:10.1007/s11618-017-0756-0
- Helmke, A. (2007). *Unterrichtsqualität erfassen, bewerten, verbessern. Dieses Buch ist Franz-Emanuel Weinert gewidmet* (Schulisches Qualitätsmanagement, 6. Aufl.). Seelze: Klett Kallmeyer.
- Helmke, A., Helmke, T., Schrader, F.-W., Wagner, W., Nold, G. & Schröder, K. (2008). Alltagspraxis des Englischunterrichts. In E. Klieme, W. Eichler, A. Helmke, R. H. Lehmann, G. Nold, H.-G. Rolff, K. Schröder, G. Thomé & H. Willenberg (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (Beltz Pädagogik, S. 371–381). Weinheim: Beltz.
- Helmke, A., Schrader, F.-W., Wagner, W., Nold, G. & Schröder, K. (2008). Selbstkonzept, Motivation und Englischleistung. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Zentrale Befunde der Studie Deutsch Englisch Schülerleistung International (DESI)* (S. 244–257). Weinheim [u.a.]: Beltz.

- Helsper, W. (1989). Jugendliche Gegenkultur und schulisch-bürokratische Rationalität. Zur Ambivalenz von Individualisierungs- und Informalisierungsprozessen. In W. Breyvogel (Hrsg.), *Pädagogische Jugendforschung. Erkenntnisse und Perspektiven* (S. 161–185). Wiesbaden: VS Verlag für Sozialwissenschaften. doi:10.1007/978-3-322-97188-3_11
- Helsper, W. & Böhme, J. (Hrsg.). (2008). *Handbuch der Schulforschung* (2. Aufl., Wiesbaden): VS Verlag für Sozialwissenschaften. doi:10.1007/978-3-531-91095-6
- Helsper, W. & Kramer, R.-T. (2001). *Schulkultur und Schulmythos. Gymnasien zwischen elitärer Bildung und höherer Volksschule im Transformationsprozeß* (Studien zur Schul- und Bildungsforschung, Bd. 13). Opladen: Leske + Budrich (Rekonstruktionen zur Schulkultur).
- Helsper, W., Sandring, S. & Wiezorek, C. (2005). Anerkennung in pädagogischen Beziehungen. Ein Problemaufriss. In W. Heitmeyer & P. Imbusch (Hrsg.), *Integrationspotenziale einer modernen Gesellschaft* (Analysen zu gesellschaftlicher Integration und Desintegration, 1. Aufl., S. 179–206). Wiesbaden: VS, Verlag für Sozialwissenschaft. doi:10.1007/978-3-322-80502-7_6
- Hertel, S., Hochweber, J., Steinert, B. & Klieme, E. (2010). Schulische Rahmenbedingungen und Lerngelegenheiten im Deutschunterricht. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, W. Schneider & P. Stanat (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 113–152). Münster: Waxmann.
- Herzog, W. (2010). Dreieck, Kreis und Stufe. Über die geometrische Selbstbegrenzung der Didaktik. In P. Bühler, T. Bühler & F. Osterwalder (Hrsg.), *Grenzen der Didaktik* (Prisma, Bd. 15, 1. Aufl., S. 157–184). Bern: Haupt.
- Hessisches Kultusministerium. (2011a). *Bildungsstandards und Inhaltsfelder. Das neue Kerncurriculum für Hessen. Sekundarstufe I – Gymnasium*. https://kultusministerium.hessen.de/sites/default/files/media/kerncurriculum_mathematik_gymnasium.pdf. Zugriffen: 15. August 2018.
- Hessisches Kultusministerium. (2011b). *Bildungsstandards und Inhaltsfelder. Das neue Kerncurriculum für Hessen. Sekundarstufe I – Realschule*. https://kultusministerium.hessen.de/sites/default/files/media/kerncurriculum_mathematik_realschule.pdf. Zugriffen: 15. August 2018.
- Hessisches Kultusministerium. (2011c). *Vom Kerncurriculum zum Schulcurriculum. Handreichung für Schulleitungen und Steuergruppen*, Wiesbaden. https://kultusministerium.hessen.de/sites/default/files/media/hkm/vom_kerncurriculum_zum_schulcurriculum_erweiterte_fassung.pdf. Zugriffen: 15. August 2018.
- Hessisches Kultusministerium. (2011d). *Verordnung über die Stundentafeln für die Primarstufe und die Sekundarstufe I. StdTafV-SekIV, HE*. http://www.rv.hessenrecht.hessen.de/lexsoft/default/hessenrecht_rv.html?doc.hl=1&doc.id=hevr-AssBFSchulAPrVHE2011rahmen&documentnumber=1&numberofresults=1&showdoccase=1&doc.part=R¶mfromHL=true.
- Hessisches Kultusministerium. (2018). *Sekundarstufe. Grundsätzliches zur Mittelstufenschule*, Hessisches Kultusministerium. <https://kultusministerium.hessen.de/schulsystem/schulwahl/schulformen/mittelstufenschule/grundsatzliches-zur-mittelstufenschule>. Zugriffen: 27. März 2018.
- Heymann, H. W. (2013). Bildungstheoretische Grundlagen, Leitbild der Gemeinschaftsschule. In T. Bohl & S. Meissner (Hrsg.), *Expertise Gemeinschaftsschule. Forschungsergebnisse und Handlungsempfehlungen für Baden-Württemberg* (Pädagogik, S. 31–45). Weinheim: Beltz.

- Hinsch, R. (1979). *Einstellungswandel und Praxisschock bei jungen Lehrern. Eine empirische Längsschnittuntersuchung* (Beltz Forschungsberichte, Dr. nach Typoskript). Weinheim: Beltz.
- Hopewell, S. (2011). Leveraging bilingualism to accelerate English reading comprehension. *International Journal of Bilingual Education and Bilingualism* 14 (5), 603–620. doi:10.1080/13670050.2011.564274
- Hopf, D. (2005). Zweisprachigkeit und Schulleistungen bei Migrantenkindern. *Zeitschrift für Pädagogik* 51 (2), 236–251.
- Horlacher, R. & Vincenti, A. de (2016). Lehrplan/Curriculum. *Enzyklopädie Erziehungswissenschaft Online*, 1–26. doi:10.3262/EEO01160362
- Hox, J. J. (2013). Multilevel regression and multilevel structural equation modeling. In T. D. Little (Hrsg.), *The Oxford handbook of quantitative methods. Statistical analysis* (Oxford Library of Psychology, Vol. 2). Oxford: Oxford University Press USA. doi:10.1093/oxfordhb/9780199934898.013.0014
- Hurrelmann, K. (2013). Das Schulsystem in Deutschland: Das „Zwei-Wege-Modell“ setzt sich durch. *Zeitschrift für Pädagogik* 59 (4), 455–468. http://www.pedocs.de/volltexte/2016/11972/pdf/ZfPaed_2013_4_Hurrelmann_Schulsystem_in_Deutschland.pdf.
- IBM Cooperation. (2013) *SPSS Statistics* [Computer software].
- Ipfing, H. J. (2009). Hauptschule. In S. Blömeke, T. Bohl, L. Haag, G. Lang-Wojtask & W. Sacher (Hrsg.), *Handbuch Schule. Theorie – Organisation – Entwicklung* (UTB, Bd. 8392, S. 236–245). Bad Heilbrunn: Klinkhardt.
- Ivanov, S. (2011a) Mathematische Kompetenz und Einstellungen zum Mathematikunterricht. In Freie und Hansestadt Hamburg – Behörde für Schule und Berufsbildung (Hrsg.), *Kompetenzen und Einstellungen von Schülerinnen und Schülern an Hamburger Schulen am Ende der Sekundarstufe I und zu Beginn der gymnasialen Oberstufe. KESS 10/11* (S. 75–122). Hamburg.
- Ivanov, S. (2011b) Naturwissenschaftliche Kompetenz und fachbezogene Einstellungen. In Freie und Hansestadt Hamburg – Behörde für Schule und Berufsbildung (Hrsg.), *Kompetenzen und Einstellungen von Schülerinnen und Schülern an Hamburger Schulen am Ende der Sekundarstufe I und zu Beginn der gymnasialen Oberstufe. KESS 10/11* (S. 187–220). Hamburg.
- Ivanov, S. & Nikolova, R. (2010). Naturwissenschaftliche Kompetenz. In W. Bos & C. Gröhlich (Hrsg.), *KESS 8. Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 8* (HANSE – Hamburger Schriften zur Qualität im Bildungswesen, Bd. 6, S. 79–99). Münster: Waxmann.
- James, M. (2006). Assessment, teaching and theories of learning. In J. Gardner (Hrsg.), *Assessment and Learning* (S. 47–60). London: Sage Publications.
- Jenßen, L., Dunekacke, S., Eid, M. & Blömeke, S. (2015). The relationship of mathematical competence and mathematics anxiety. *Zeitschrift für Psychologie* 223 (1), 31–38. doi:10.1027/2151-2604/a000197
- Jose, P. E. (2013). *Doing Statistical Mediation and Moderation* (Methodology in the social sciences). New York: Guilford Publications Inc.
- Kaas, R. E. & Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association* 90 (430), 773–795. doi:10.1080/01621459.1995.10476572
- Kant, I. (1784). Beantwortung der Frage: Was ist Aufklärung? *Berlinische Monatsschrift Dezember-Heft*, 481–494.

- Keck, R. W. (2009). Zur Geschichte der Schule. In S. Blömeke, T. Bohl, L. Haag, G. Lang-Wojtask & W. Sacher (Hrsg.), *Handbuch Schule. Theorie – Organisation – Entwicklung* (UTB, Bd. 8392, S. 157–162). Bad Heilbrunn: Klinkhardt.
- Kingston, N. & Nash, B. (2011). Formative Assessment. A Meta-Analysis and a Call for Research. *Educational Measurement: Issues and Practice* 30 (4), 28–37. doi:10.1111/j.1745-3992.2011.00220.x
- Kiper, H. (2007a). Pädagogik des Gymnasiums – quo vadis? In S. Jahnke-Klein, H. Kiper & L. Freisel (Hrsg.), *Gymnasium heute. Zwischen Elitebildung und Förderung der Vielen* (S. 37–50). Baltmannsweiler: Schneider Verl. Hohengehren.
- Kiper, H. (2007b). Veränderungsprozesse im Gymnasium heute. Zur Bedeutung neuer Steuerungsinstrumente. In S. Jahnke-Klein, H. Kiper & L. Freisel (Hrsg.), *Gymnasium heute. Zwischen Elitebildung und Förderung der Vielen* (S. 69–92). Baltmannsweiler: Schneider Verl. Hohengehren.
- Kiper, H. & Mischke, W. (2004). *Einführung in die Allgemeine Didaktik* (Beltz Studium). Weinheim: Beltz.
- Klafki, W. (1976). Handlungsforschung im Schulfeld. In W. Klafki (Hrsg.), *Aspekte kritisch-konstruktiver Erziehungswissenschaft. Gesammelte Beiträge zur Theorie-Praxis-Diskussion* (Beltz-Studienbuch, Bd. 104, S. 59–96). Weinheim: Beltz.
- Klafki, W. (1996). *Neue Studien zur Bildungstheorie und Didaktik. Zeitgemäße Allgemeinbildung und kritisch-konstruktive Didaktik* (Reihe Pädagogik, 5., unveränd. Aufl.). Weinheim: Beltz.
- Klieme, E. (2006). Empirische Unterrichtsforschung. Aktuelle Entwicklungen, theoretische Grundlagen und fachspezifische Befunde. *Zeitschrift für Pädagogik* 52 (6), 765–773.
- Klieme, E. (2018). Unterrichtsqualität. In M. Gläser-Zikuda, M. Harring & C. Rohlf (Hrsg.), *Handbuch Schulpädagogik* (1. Aufl.,). Münster: Waxmann, UTB.
- Klieme, E., Eichler, W., Helmke, A., Lehmann, R. H., Nold, G., Rolff, H.-G., Schröder K., Thomé, G. & Willenberg, H. (Hrsg.). (2008). *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (Beltz Pädagogik). Weinheim: Beltz.
- Klieme, E., Jude, N., Rauch, D., Ehlers, H., Helmke, A., Eichler, W., Nold, G., Thomé, G. & Willenberg, H. (2008). Alltagspraxis, Qualität und Wirksamkeit des Deutschunterrichts. In E. Klieme, W. Eichler, A. Helmke, R. H. Lehmann, G. Nold, H.-G. Rolff, K. Schröder, G. Thomé & H. Willenberg (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (Beltz Pädagogik, S. 319–355). Weinheim: Beltz.
- Klieme, E., Neubrand, M. & Lüdtke, O. (2001). Mathematische Grundbildung. Testkonzeption und Ergebnisse. In Deutsches PISA Konsortium (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 139–190). Opladen: Leske + Budrich. doi:10.1007/978-3-322-83412-6_5
- Klieme, E., Pauli, C. & Reusser, K. (2009). The Pythagoras Study. Investigating effects of teaching and learning in Swiss and German mathematics classrooms. In T. Janik & T. Seidel (Hrsg.), *The power of video studies in investigating teaching and learning in the classroom* (S. 137–160). Münster: Waxmann.
- Klieme, E. & Rakoczy, K. (2001). Unterrichtsqualität aus Schülerperspektive. Kulturspezifische Profile, regionale Unterschiede und Zusammenhänge mit Effekten von Unterricht. In Deutsches PISA Konsortium (Hrsg.), *PISA 2000. Basiskompetenzen*

- von Schülerinnen und Schülern im internationalen Vergleich (S. 333–359). Opladen: Leske + Budrich. doi:10.1007/978-3-322-97590-4_12
- Klieme, E., Schümer, G. & Knoll, S. (2001). Mathematikunterricht in der Sekundarstufe I. „Aufgabenkultur“ und Unterrichtsgestaltung. In Bundesministerium für Bildung und Forschung (Hrsg.), *TIMSS. Impulse für Schule und Unterricht*. Forschungsbe-funde, Reforminitiativen, Praxisberichte und Video-Dokumente (S. 43–57). Mün-chen: Mediahaus Biering Grafischer Betrieb.
- Klieme, E., Steinert, B. & Hochweber, J. (2010). Zur Bedeutung der Schulqualität für Unterricht und Lernergebnisse. In W. Bos (Hrsg.), *Schulische Lerngelegenheiten und Kompetenzentwicklung. Festschrift für Jürgen Baumert* (S. 231–255). Münster: Waxmann.
- Klinge, D. (2016). *Die elterliche Übergangsentscheidung nach der Grundschule. Werte, Erwartungen und Orientierungen* (1. Aufl.). Wiesbaden: Springer Fachmedien; Springer VS. doi:10.1007/978-3-658-14351-0
- Klingner, J. K. & Vaughn, S. (2000). The helping behaviors of fifth graders while using collaborative strategic reading during ESL content classes. *TESOL Quarterly* 34 (1), 69. doi:10.2307/3588097
- Klusmann, U. (2011). Allgemeine berufliche Motivation und Selbstregulation. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Pro-fessionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 277–294). Münster: Waxmann.
- Klusmann, U., Kunter, M., Trautwein, U., Lüdtke, O. & Baumert, J. (2008). Teachers’ occupational well-being and quality of instruction. The important role of self-regulatory patterns. *Journal of Educational Psychology* 100 (3), 702–715. doi:10.1037/0022-0663.100.3.702
- KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bun-desrepublik Deutschland) (1994). *Empfehlungen zur sonderpädagogischen Förde-rung in den Schulen in der Bundesrepublik Deutschland*. <https://www.kmk.org/fileadmin/pdf/PresseUndAktuelles/2000/sopae94.pdf>. Zugriffen: 16. Juni 2018.
- KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bun-desrepublik Deutschland, Hrsg.). (1997). *Rahmenvereinbarung über die Ausbildung und Prüfung für übergreifende Lehrämter der Primarstufe und aller oder einzelner Schularten der Sekundarstufe I (Lehramtstyp 2)*. https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/1997/1997_02_28-RV_Lehramtstyp_2.pdf. Zugriffen: 26. Juli 2018.
- KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bun-desrepublik Deutschland) (2003). *Vereinbarung über Bildungsstandards für den Mittleren Abschluss (Jahrgangsstufe 10). Beschluss der Kultusministerkonferenz vom 04.12.2003*. https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2003/2003_12_04-Vereinbarung-Bildungsstandards-MS.pdf. Zugriffen: 15. August 2018.
- KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bun-desrepublik Deutschland) (2006). *Vereinbarung über die Schularten und Bildungs-gänge im Sekundarbereich I*. Beschluss der Kultusministerkonferenz vom 03.12.1993 i.d.F. vom 02.06.2006. [https:// www.google.de/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwj67Kr4-srZAhXHY1AKHUOVcpGQFggnMAA&url=http%3A%2F%2Fwww.kmk.org%2Ffileadmin%2Fpdf%2FPresseUndAktuelles%2FBeschluesse_Veroeffentlichungen%2Fallg_Schulwes](https://www.google.de/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwj67Kr4-srZAhXHY1AKHUOVcpGQFggnMAA&url=http%3A%2F%2Fwww.kmk.org%2Ffileadmin%2Fpdf%2FPresseUndAktuelles%2FBeschluesse_Veroeffentlichungen%2Fallg_Schulwes)

- en%2FSchulart_Bildungsg_Sek1.pdf&usg=AOvVaw3X6XiGjnyHOibiDVf5mMdp. Zugriffen: 1. März 2018.
- KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, Hrsg.). (2014). *Vereinbarungen über die Schularten und Bildungsgänge im Sekundarbereich I. Beschluss der Kultusministerkonferenz vom 03.12.1993 i.d.F. vom 25.09.2014*. https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/1993/1993_12_03-VB-Sek-I.pdf. Zugriffen: 26. Februar 2018.
- KMK. (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, Hrsg.). (2016a). *Sonderpädagogische Förderung in Förderschulen (Sonderschulen)*. Berlin. https://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Dokumentationen/Aus_SoPae_2015.pdf. Zugriffen: 19. März 2018.
- KMK. (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, Hrsg.). (2016b). *Sonderpädagogische Förderung in Schulen 2005 bis 2014* (Statistische Veröffentlichungen der Kultusminister Konferenz Dokumentation Nr. 210). https://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Dokumentationen/Dok_210_SoPae_2014.pdf. Zugriffen: 16. März 2018.
- KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, Hrsg.). (2017a). *Das Bildungswesen in der Bundesrepublik Deutschland 2014/2015. Darstellung der Kompetenzen, Strukturen und bildungspolitischen Entwicklungen für den Informationsaustausch in Europa*, Bonn. https://www.kmk.org/fileadmin/Dateien/pdf/Eurydice/Bildungswesen-dt-pdfs/dossier_de_ebook.pdf.
- KMK. (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, Hrsg.). (2017b). *Definitionenkatalog zur Schulstatistik 2017*. Bonn. <https://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Defkat2017.pdf>. Zugriffen: 19. März 2018.
- KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, Hrsg.). (2017c). *Sachstand der Lehrerbildung*. https://www.kmk.org/fileadmin/Dateien/pdf/Bildung/AllgBildung/2017-03-07_Sachstand_LB_o_EW.pdf. Zugriffen: 26. Juli 2018.
- KMK. (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, Hrsg.). (2018). *Lehrplandatenbank*. <https://www.kmk.org/dokumentation-statistik/rechtsvorschriften-lehrplaene/lehrplan-datenbank.html>. Zugriffen: 15. August 2018.
- Knab, D. (1983). Der Beitrag der Curriculumforschung zu Erziehungswissenschaft und Bildungstheorie. Versuch einer Zwischenbilanz. In U. Hameyer, K. Frey, H. Haft & Hameyer (Hrsg.), *Handbuch der Curriculumforschung. Erste Ausgabe – Übersichten zur Forschung 1970–1981* (S. 697–711). Weinheim: Beltz.
- Knobloch, P. D. (2014). Internationalisierung als Herausforderung für die Pädagogik. Entwurf eines meta-theoretischen Vermittlungsansatzes zwischen Vergleichender Erziehungswissenschaft und Allgemeiner Pädagogik 60 (5), 724–735.
- Kohlberg, L. (2014). *Die Psychologie der Moralentwicklung* (Beiträge zur Soziogenese der Handlungsfähigkeit, Bd. 1232, 7. Aufl.). Frankfurt am Main: Suhrkamp.
- Köhler, S.-M. (2016). Die sozialisationstheoretische Perspektive. Der Wandel der Peer- und Freundschaftsbeziehungen im Lebensverlauf. In S.-M. Köhler, H.-H. Krüger & N. Pfaff (Hrsg.), *Handbuch Peerforschung* (89-119). Opladen: Verlag Barbara Budrich. doi:10.2307/j.ctvd7w8m2.8

- Köhler, S.-M., Krüger, H.-H. & Pfaff, N. (2016). Peergroups als Forschungsgegenstand. Einleitung. In S.-M. Köhler, H.-H. Krüger & N. Pfaff (Hrsg.), *Handbuch Peerforschung* (S. 11–33). Opladen: Verlag Barbara Budrich. doi:10.2307/j.ctvd7w8m2.4
- Köller, O. (1996). Die Entwicklung der Schulleistungen und psychosozialer Merkmale während der Sekundarstufe. In Max-Planck-Institut für Bildungsforschung (Hrsg.), *Bildungsverläufe und psychosoziale Entwicklung im Jugendalter (BIJU). 2. Bericht für die Schulen* (S. 13–24). Berlin: Max-Planck-Institut für Bildungsforschung.
- Köller, O. (2004). *Konsequenzen von Leistungsgruppierungen* (Pädagogische Psychologie und Entwicklungspsychologie, Bd. 37). Münster: Waxmann.
- Köller, O. (2007). Das Gymnasium zwischen Elitebildung und Förderung der Vielen. Welche Pädagogik braucht das Gymnasium. In S. Jahnke-Klein, H. Kiper & L. Freisel (Hrsg.), *Gymnasium heute. Zwischen Elitebildung und Förderung der Vielen* (S. 13–35). Baltmannsweiler: Schneider Verl. Hohengehren.
- Köller, O. (2008). Gesamtschule. Erweiterung statt Alternative. In K. S. Cortina, J. Baumert, A. Leschinsky, K. U. Mayer & L. Trommer (Hrsg.), *Das Bildungswesen in der Bundesrepublik Deutschland* (S. 437–465). Reinbek bei Hamburg: Rowohlt-Taschenbuch-Verlag.
- Köller, O. & Baumert, J. (2001). Leistungsgruppierung in der Sekundarstufe I. Ihre Konsequenzen für die Mathematikleistung und das mathematische Selbstkonzept der Begabung. *Zeitschrift für Pädagogische Psychologie* 15 (2), 99–110. doi:10.1024//1010-0652.15.2.99
- Köller, O., Baumert, J. & Schnabel, K. (2000). Zum Zusammenspiel von schulischem Interesse und Lernen im Fach Mathematik. Längsschnittanalysen in den Sekundarstufen I und II. In U. Schiefele (Hrsg.), *Interesse und Lernmotivation. Neue Studien zu Entwicklung und Wirkungen* (S. 163–182). Münster: Waxmann.
- Konrad, F.-M. (2010). *Wilhelm von Humboldt*. Bern: Haupt, UTB.
- Kounin, J. S. (1970). *Discipline and group management in classrooms*. New York, NY: Holt Rinehart & Winston.
- Krappmann, L. (2004). Sozialisation in Interaktionen und Beziehungen unter Gleichaltrigen in der Schulklasse. In D. Geulen & H. Veith (Hrsg.), *Sozialisationstheorie interdisziplinär. Aktuelle Perspektiven* (Der Mensch als soziales und personales Wesen, Bd. 20, Reprint 2016, S. 253–271). Berlin: De Gruyter Oldenbourg. doi:10.1515/9783110511246-015
- Krimm, S. (2009). Die höheren Sphären und die Realien. Überlegungen zu Geschichte und Perspektiven gymnasialer Bildung. In D. Bosse (Hrsg.), *Gymnasiale Bildung zwischen Kompetenzorientierung und Kulturarbeit* (1. Aufl., S. 242–258). Wiesbaden: VS Verl. für Sozialwiss. doi:10.1007/978-3-531-91485-5_18
- Kron, F. W., Jürgens, E. & Standop, J. (2014). *Grundwissen Didaktik* (utb-studi-e-book, 6. überarb. Aufl.). München: Ernst Reinhardt Verlag.
- Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology* 5 (3), 213–236. doi:10.1002/acp.2350050305
- Krosnick, J. A. (2000). The threat of satisficing in surveys. The shortcuts respondents take in answering questions. *Survey Methods Newsletter* 20 (1), 4–8.
- Kuger, S. (2016). Curriculum and learning time in international school achievement studies. In S. Kuger, E. Klieme, N. Jude & D. Kaplan (Hrsg.), *Assessing contexts of learning. An international perspective* (Methodology of educational measurement

- and assessment, S. 395–422). Cham: Springer International Publishing. doi:10.1007/978-3-319-45357-6_16
- Kuger, S., Klieme, E., Lüdtke, O., Schiepe-Tiska, A. & Reiss, K. (2017). Mathematikunterricht und Schülerleistung in der Sekundarstufe. Zur Validität von Schülerbefragungen in Schulleistungsstudien. *Zeitschrift für Erziehungswissenschaft* 20 (S2), 61–98. doi:10.1007/s11618-017-0750-6
- Kuhlmann, C. (2013). *Erziehung und Bildung. Einführung in die Geschichte und Aktualität pädagogischer Theorien* (Lehrbuch). Wiesbaden: Springer VS.
- Kunter, M. (2005). *Multiple Ziele im Mathematikunterricht* (Pädagogische Psychologie und Entwicklungspsychologie, Bd. 51). Münster: Waxmann.
- Kunter, M. & Trautwein, U. (2013). *Psychologie des Unterrichts* (StandardWissen Lehramt, Bd. 3895). Paderborn: Ferdinand Schöningh.
- Kunter, M. & Voss, T. (2011). Das Modell der Unterrichtsqualität in COACTIV. Eine multikriteriale Analyse. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 85–113). Münster: Waxmann.
- Kunter, M., Brunner, M., Baumert, J., Klusmann, U., Krauss, S., Blum, W., Jordan, A. & Neubrand, M. (2005). Der Mathematikunterricht der PISA-Schülerinnen und -Schüler. *Zeitschrift für Erziehungswissenschaft* 8 (4), 502–520. doi:10.1007/s11618-005-0156-8
- Kunter, M., Kleickmann, T., Klusmann, U. & Richter, D. (2011). Die Entwicklung professioneller Kompetenz. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 55–68). Münster: Waxmann.
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (Hrsg.). (2011). *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*. Münster: Waxmann.
- Künzli, R. (2009). Curriculum und Lehrmittel. In S. Andresen, R. Casale, T. Gabriel, R. Horlacher, S. Larcher Klee & J. Oelkers (Hrsg.), *Handwörterbuch Erziehungswissenschaft* (Studium Paedagogik, S. 134–148). Weinheim: Beltz.
- Kurz, A. (2011). Access to what should be taught and will be tested. Students' opportunity to learn the intended curriculum. In S. N. Elliott, R. J. Kettler, P. A. Beddow & A. Kurz (Hrsg.), *The handbook of accessible achievement tests for all students. Bridging the gaps between research, practice, and policy*. Springer: New York; NY. doi:10.1007/978-1-4419-9356-4_6
- Lang-Wojtasik, G., Kansteiner, K. & Stratmann, J. (2016). Gemeinschaftsschule als pädagogische und gesellschaftliche Herausforderung. In G. Lang-Wojtasik, K. Kansteiner & J. Stratmann (Hrsg.), *Gemeinschaftsschule als pädagogische und gesellschaftliche Herausforderung* (S. 7–16). Münster: Waxmann.
- Lechner, C. M. & Rammstedt, B. (2015). Cognitive ability, acquiescence, and the structure of personality in a sample of older adults. *Psychological assessment* 27 (4), 1301–1311. doi:10.1037/pas0000151
- Leschinsky, A. (2008a). Die Hauptschule. von der Be- zur Enthauptung. In K. S. Cortina, J. Baumert, A. Leschinsky, K. U. Mayer & L. Trommer (Hrsg.), *Das Bildungswesen in der Bundesrepublik Deutschland* (S. 377–406). Reinbek bei Hamburg: Rowohlt-Taschenbuch-Verlag.
- Leschinsky, A. (2008b). Die Realschule. Ein zweischneidiger Erfolg. In K. S. Cortina, J. Baumert, A. Leschinsky, K. U. Mayer & L. Trommer (Hrsg.), *Das Bildungswesen in*

- der Bundesrepublik Deutschland (S. 407–435). Reinbek bei Hamburg: Rowohlt-Taschenbuch-Verlag.
- Leschinsky, A. & Cortina, K. S. (2008). Zur sozialen Einbettung bildungspolitischer Trends in der Bundesrepublik. In K. S. Cortina, J. Baumert, A. Leschinsky, K. U. Mayer & L. Trommer (Hrsg.), *Das Bildungswesen in der Bundesrepublik Deutschland* (S. 21–51). Reinbek bei Hamburg: Rowohlt-Taschenbuch-Verlag.
- Liebau, E., Mack, W. & Scheilke, C. T. (1997). Einleitung. Das Gymnasium im Wandel. In E. Liebau, W. Mack & C. T. Scheilke (Hrsg.), *Das Gymnasium. Alltag, Reform, Geschichte, Theorie* (Grundlagentexte Pädagogik, S. 7–14). Weinheim: Juventa.
- Liegmann, A. B. (2016). Die Ordnungen der Schulformen. Ein bundesweiter Vergleich der Schulformen mit mehreren Bildungsgängen. In T.-S. Idel, F. Dietrich, K. Kunze, K. Rabenstein & A. Schütz (Hrsg.), *Professionsentwicklung und Schulstrukturenform. Zwischen Gymnasium und neuen Schulformen in der Sekundarstufe* (Studien zur Professionsforschung und Lehrerbildung, S. 47–63). Bad Heilbrunn: Klinkhardt.
- Lipowsky, F. (2015). Unterricht. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 69–105). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-41291-2_4
- Liu, M., Harbaugh, A. G., Harring, J. R. & Hancock, G. R. (2017). The effects of extreme response and non-extreme response styles on testing measurement invariance. *Frontiers in Psychology* 8. doi:10.3389/fpsyg.2017.00726
- Lubke, G. H. & Muthén, B. O. (2004). Applying multigroup confirmatory factor models for continuous outcomes to likert scale data complicates meaningful group comparisons. *Structural Equation Modeling* 11 (4), 514–534. doi:10.1207/s15328007sem1104_2
- Lüders, M. (2012). Der Unterrichtsbegriff in pädagogischen Nachschlagewerken. *Zeitschrift für Pädagogik* 58 (1), 109–129.
- Maaz, K. (2017). *Chancengerechtigkeit im Deutschen Bildungssystem* (DUK-Publikationsreihe „Blickwinkel“). Berlin: Druckerei Brandt.
- Maaz, K., Baumert, J., Gresch, C. & McElvany, N. (2010). *Der Übergang von der Grundschule in die weiterführende Schule. Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten* (Bildungsforschung, Bd. 34). Berlin: Bundesministerium für Bildung und Forschung (BMBF).
- Maaz, K., Baumert, J., Neumann, M., Becker, M. & Dumont, H. (Hrsg.). (2013). *Die Berliner Schulstrukturreform. Bewertung durch die beteiligten Akteure und Konsequenzen des neuen Übergangsverfahrens von der Grundschule in die weiterführenden Schulen*. Münster: Waxmann.
- MacKinnon, D. (2012). *Introduction to statistical mediation analysis*. Hoboken: Taylor and Francis. doi:10.1002/9781118133880.hop202025
- Maestri, V. (2011). *Can ethnic diversity have a positive effect on school achievement?* https://mpra.ub.uni-muenchen.de/33455/1/MPRA_paper_33455.pdf. Zugriffen: 2. August 2018.
- Malecki, A. (2016). *Schulen auf einen Blick. Ausgabe 2016* (Statistisches Bundesamt, Hrsg.), Wiesbaden. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bildung-Forschung-Kultur/Schulen/Publikationen/Downloads-Schulen/broschuere-schulen-blick-0110018169004.pdf?__blob=publicationFile Zugriffen: 22. März 2018.
- Mang, J., Ustjanzew, N., Schiepe-Tiska, A., Prenzel, M., Sälzer, C., Müller, K. & González Rodríguez, E. (2018). *PISA 2012 Skalenhandbuch. Dokumentation der Erhebungsinstrumente*. Münster: Waxmann.

- Manski, C. F. (2000). Economic analysis of social interactions. *Journal of Economic Perspectives* 14 (3), 115–136. doi:10.1257/jep.14.3.115
- Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of Educational Psychology* 79 (3), 280–295. doi:10.1037/0022-0663.79.3.280
- Marsh, H. W. (1991). Failure of high-ability high schools to deliver academic benefits commensurate with their students' ability levels. *American Educational Research Journal* 28 (2), 445–480. doi:10.3102/00028312028002445
- Marsh, H. W. (2005). Big-fish-little-pond effect on academic self-concept. *Zeitschrift für Pädagogische Psychologie* 19 (3), 119–129. doi:10.1024/1010-0652.19.3.119
- Marsh, H. W. & O'Mara, A. J. (2010). Long-term total negative effects of school-average ability on diverse educational outcomes. *Zeitschrift für Pädagogische Psychologie* 24 (1), 51–72. doi:10.1024/1010-0652/a000004
- Marsh, H. W., Trautwein, U., Lüdtke, O., Baumert, J. & Köller, O. (2016). The big-fish-little-pond effect. Persistent negative effects of selective high schools on self-concept after graduation. *American Educational Research Journal* 44 (3), 631–669. doi:10.3102/0002831207306728
- Marsh, H. W., Trautwein, U., Lüdtke, O. & Köller, O. (2008). Social comparison and big-fish-little-pond effects on self-concept and other self-belief constructs. Role of generalized and specific others. *Journal of Educational Psychology* 100 (3), 510–524. doi:10.1037/0022-0663.100.3.510
- Martens, M. (2010). Schulformunterschiede im Umgang mit Darstellungen von Geschichte. Hauptschule, Realschule und Gymnasium im Vergleich. *Zeitschrift für Geschichtsdidaktik* 9, 57–78. doi:10.13109/zfgd.2010.09.1.57
- Maskus, R. (1966). *Zur Geschichte der Mittel- und Realschule*. Bad Heilbrunn: Klinkhardt.
- Matsunaga, M. (2010). How to factor-analyze your data right. Do's and don'ts and how-to's. *International Journal of Psychological Research* 3 (1), 97–110. doi:10.21500/20112084.854
- Matthiessen, W. (2000). *Geschichte* (Pocket Teacher Abi, 1. Aufl., 1. Dr.). Berlin: Cornelsen Scriptor.
- Max-Planck-Institut für Bildungsforschung (Hrsg.). (1996). *Bildungsverläufe und psychosoziale Entwicklung im Jugendalter (BIJU). 2. Bericht für die Schulen*. Berlin: Max-Planck-Institut für Bildungsforschung.
- May, P. (2010). Orthografische Kompetenz. In W. Bos & C. Gröhlich (Hrsg.), *KESS 8. Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 8* (HANSE – Hamburger Schriften zur Qualität im Bildungswesen, Bd. 6, S. 67–78). Münster: Waxmann.
- Mead, G. H. (1968). *Geist, Identität und Gesellschaft aus der Sicht des Sozialbehaviorismus*. Frankfurt am Main: Suhrkamp.
- Meisenberg, G. & Williams, A. (2008). Are acquiescent and extreme response styles related to low intelligence and education? *Personality and Individual Differences* 44 (7), 1539–1550. doi:10.1016/j.paid.2008.01.010
- Meseth, W. & Proske, M. (2011). Unterricht. In J. Kade (Hrsg.), *Pädagogisches Wissen. Erziehungswissenschaft in Grundbegriffen* (Grundriss der Pädagogik, Erziehungswissenschaft, Bd. 5, S. 102–108). Stuttgart: Kohlhammer.
- Meseth, W., Proske, M. & Radtke, F.-O. (2011). Ein Streitgespräch zwischen Andreas Gruschka, Walter Herzog, Wolfgang Meseth, Matthias Proske und Sabine Reh. „In der Frage der Unterrichtstheorien stehen wir doch ziemlich am Anfang“. In W. Me-

- seth, M. Proske & F.-O. Radtke (Hrsg.), *Unterrichtstheorien in Forschung und Lehre* (S. 242–262). Bad Heilbrunn: Verlag Julius Klinkhardt.
- Meyer, A. (2010). *Die Epoche der Aufklärung* (Akademie Studienbücher 6-2010). München: Oldenbourg Akademieverlag. doi:10.1524/9783050049496
- Meyer, H. (2002). Unterrichtsmethoden. In H. Kiper, H. Meyer, W. Topsch & R. Hinz (Hrsg.), *Einführung in die Schulpädagogik* (Studium kompakt. Unterricht, Schule, S. 109–121). Berlin: Cornelsen.
- Meyer, H. (2017). *Was ist guter Unterricht?* (12. Auflage). Berlin: Cornelsen.
- Miller, S. (2013). Methodische Vielfalt für heterogene Lerngruppen. In A. Kaiser, D. Schmetz, P. Wachtel & B. Werner (Hrsg.), *Didaktik und Unterricht* (Behinderung, Bildung, Partizipation, Bd. 4, S. 146–156). Stuttgart: Kohlhammer.
- Miller, G. M. & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology* 110 (1), 40–48. doi:10.1037//0021-843X.110.1.40
- Millsap, R. E. (2011). *Statistical approaches to measurement invariance*. New York: Psychology Press. doi:10.4324/9780203821961
- Moosbrugger, H. (1978). *Multivariate statistische Analyseverfahren* (Kohlhammer Standards Psychologie : Studententext : Teilgebiet Methoden, 1. Aufl.). Stuttgart: Kohlhammer.
- Muthén, B. & Asparouhov, T. (2014). IRT studies of many groups. The alignment method. *Frontiers in Psychology* 5. doi:10.3389/fpsyg.2014.00978
- Muthén, L. K. & Muthén, B. O. (1998–2013) *Mplus7* [Computer software].
- Muthén, L. K. & Muthén, B. O. (2017). *Mplus User's Guide* (8th). Los Angeles, CA: Muthén & Muthén.
- Nash, R. (2003). Inequality/difference in education. Is a real explanation of primary and secondary effects possible? *The British journal of sociology* 54 (4), 433–451. doi:10.1080/0007131032000143537
- Nath, A. (2001). Die Perioden des modernen Bildungswachstums. In H. J. Apel (Hrsg.), *Das öffentliche Bildungswesen. Historische Entwicklung, gesellschaftliche Funktionen, pädagogischer Streit* (S. 14–48). Bad Heilbrunn/Obb.: Klinkhardt.
- National Board for Professional Teaching Standards. (2002). *What teachers should know and be able to do* (NBPTS, Hrsg.), Arlington, VA. <http://accomplishedteacher.org/wp-content/uploads/2016/12/NBPTS-What-Teachers-Should-Know-and-Be-Able-to-Do-.pdf>. Zugriffen: 26. Juli 2018.
- Naumann, A., Musow, S., Aichele, C., Hochweber, J. & Hartig, J. (2019). Instruktions-sensitivität von Tests und Items. *Zeitschrift für Erziehungswissenschaft* 22, 181–202. doi:10.1007/s11618-018-0832-0
- Naumann, J., Artelt, C., Schneider, W. & Stanat, P. (2013). Lesekompetenz in PISA 2000 bis 2009. In N. Jude & E. Klieme (Hrsg.), *PISA 2009. Impulse für die Schul- und Unterrichtsforschung* (S. 23–71). Weinheim: Beltz.
- Neubrand, M., Jordan, A., Krauss, S., Blum, W. & Löwen, K. (2011). Aufgaben im COACTIV-Projekt: Einblicke in das Potenzial für kognitive Aktivierung im Mathematikunterricht. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 115–131). Münster: Waxmann.
- Nikolova, R. (2011a) Englischleistungen und Einstellungen zum Englischunterricht. In Freie und Hansestadt Hamburg – Behörde für Schule und Berufsbildung (Hrsg.), *Kompetenzen und Einstellungen von Schülerinnen und Schülern an Hamburger*

- Schulen am Ende der Sekundarstufe I und zu Beginn der gymnasialen Oberstufe. KESS 10/11* (S. 123–160). Hamburg.
- Nikolova, R. (2011b) Lesekompetenz und Einstellungen zum Deutschunterricht. In Freie und Hansestadt Hamburg – Behörde für Schule und Berufsbildung (Hrsg.), *Kompetenzen und Einstellungen von Schülerinnen und Schülern an Hamburger Schulen am Ende der Sekundarstufe I und zu Beginn der gymnasialen Oberstufe. KESS 10/11* (S. 31–74). Hamburg.
- Nikolova, R. & Ivanov, S. (2010). Englischleistungen. In W. Bos & C. Gröhlich (Hrsg.), *KESS 8. Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 8* (HANSE – Hamburger Schriften zur Qualität im Bildungswesen, Bd. 6, S. 49–66). Münster: Waxmann.
- Noack, M. (2007). *Faktorenanalyse* (Universität Duisburg-Essen, Hrsg.). <https://www.uni-due.de/imperia/md/content/soziologie/stein/faktorenanalyse.pdf>. Zugegriffen: 24. Oktober 2017.
- Nold, D. (2010). Sozioökonomischer Status von Schülerinnen und Schülern 2008. *Wirtschaft und Statistik* 2, 138–149.
- OECD. (2010). *Mathematics teaching and learning strategies in PISA*. Paris: OECD.
- OECD. (2013). *What do students think about school* (PISA in Focus Nr. 24). <https://doi.org/10.1787/22260919>. Zugegriffen: 28. August 2018.
- OECD. (2014). *PISA 2012. Technical Report*. <https://www.oecd.org/pisa/pisaproducts/PISA-2012-technical-report-final.pdf> Zugegriffen: 10. September 2018.
- OECD. (2017). *Education at a Glance 2017*. Paris: OECD Publishing.
- Op't Eynde, P., Corte, E. de & Verschaffel, L. (2003). Framing students' mathematics-related beliefs. A quest for conceptual clarity and a comprehensive categorization. In G. C. Leder, E. Pehkonen & G. Törner (Hrsg.), *Beliefs. A hidden variable in mathematics education?* (Mathematics Education Library, Bd. 31, S. 13–36). Dordrecht: Kluwer Academic Publishers. doi:10.1007/0-306-47958-3_2
- Oser, F. (1981). *Moralisches Urteil in Gruppen, soziales Handeln, Verteilungsgerechtigkeit. Stufen der interaktiven Entwicklung und ihre erzieherische Stimulation* (Suhrkamp Taschenbuch Wissenschaft, Bd. 335). Frankfurt am Main: Suhrkamp.
- Pajares, M. F. (1992). Teachers' beliefs and educational research. Cleaning up a messy construct. *Review of Educational Research* 62, 307–332. doi:10.3102/00346543062003307
- Parsons, T. (1949). *The structure of social action*. Glencoe: Free Press.
- Peisert, H. (1967). *Soziale Lage und Bildungschancen in Deutschland*. München: Piper.
- Pekrun, R. & Zirngibl, A. (2004). Schülermerkmale im Fach Mathematik. In PISA-Konsortium (Hrsg.), *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland. Ergebnisse des zweiten internationalen Vergleichs* (S. 191–210). Münster: Waxmann.
- Pfost, M., Karing, Constance, Lorenz, C., Artelt, C. & Karing, C. (2010). Schereneffekte im ein- und mehrgliedrigen Schulsystem. Differenzielle Entwicklung sprachlicher Kompetenz am Übergang von der Grund in die weiterführende Schule. *Zeitschrift für Pädagogische Psychologie* 24 (3–4), 259–272. doi:10.1024/1010-0652/a000025
- Piaget, J. (1977). *The development of thought. Equilibration of cognitive structures*. New York: The Viking Press.
- Piaget, J. (1986). *Das moralische Urteil beim Kinde* (Ungekürzte Ausgabe). München: Deutscher Taschenbuch Verlag.

- Pinar, W. F. (2009). Curriculumentwicklung. In S. Andresen, R. Casale, T. Gabriel, R. Horlacher, S. Larcher Klee & J. Oelkers (Hrsg.), *Handwörterbuch Erziehungswissenschaft* (Studium Paedagogik, S. 149–162). Weinheim: Beltz.
- PISA-Konsortium (Hrsg.). (2004). *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland. Ergebnisse des zweiten internationalen Vergleichs*. Münster: Waxmann.
- Praetorius, A.-K., Klieme, E., Herbert, B. & Pinger, P. (2018). Generic dimensions of teaching quality. The German framework of Three Basic Dimensions. *ZDM* 50 (3), 407–426. doi:10.1007/s11858-018-0918-4
- Prange, K. (2011). Didaktik und Methodik. In J. Kade (Hrsg.), *Pädagogisches Wissen. Erziehungswissenschaft in Grundbegriffen* (Grundriss der Pädagogik, Erziehungswissenschaft, Bd. 5, S. 183–188). Stuttgart: Kohlhammer.
- Prenzel, M., Sälzer, C., Klieme, E. & Köller, O. (Hrsg.). (2013). *PISA 2012. Fortschritte und Herausforderungen in Deutschland*. Münster: Waxmann.
- Preuss-Lausitz, U. (2008). Gemeinschaftsschule als Antwort auf die Krise der Schule? In U. Preuss-Lausitz (Hrsg.), *Gemeinschaftsschule – Ausweg aus der Schulkrise? Konzepte, Erfahrungen, Problemlösungen* (Bibliothek Beltz, 1. Aufl., S. 9–24). Weinheim: Beltz.
- Preuss-Lausitz, U. (2018). Separation oder Inklusion. Zur Entwicklung der sonderpädagogischen Förderung im Kontext der allgemeinen Schulentwicklung. In F. J. Müller, A. Sander, H. Eberwein, H. Reiser, J. Schöler, R. Maikowski, R. Kornmann, U. Preuss-Lausitz, U. Schildmann & W. Jantzen (Hrsg.), *Blick zurück nach vorn – WegbereiterInnen der Inklusion: Band 1* (Dialektik der Be-Hinderung, S. 245–267). Gießen: Psychosozial-Verlag.
- Protsch, P. (2013). Höhere Anforderungen in der beruflichen Erstausbildung? *WSI Mitteilungen* (1), 15–22. doi:10.5771/0342-300X-2013-1-15
- Protz, S. (2004). Unterricht. In D. Benner (Hrsg.), *Historisches Wörterbuch der Pädagogik* (S. 1031–1070). Weinheim: Beltz.
- Raithel, J., Dollinger, B. & Hörmann, G. (2007). *Einführung Pädagogik. Begriffe · Strömungen Klassiker · Fachrichtungen* (2., durchgesehene und erweiterte Auflage). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Rakoczy, K., Klieme, E., Bürgermeister, A. & Harks, B. (2008). The interplay between student evaluation and instruction. *Zeitschrift für Psychologie / Journal of Psychology* 216 (2), 111–124. doi:10.1027/0044-3409.216.2.111
- Rasch, B., Friese, M., Hofmann, W. & Naumann, E. (2008). *Quantitative Methoden* (Bd. 2, 2. Auflage, 2 Bände). Heidelberg: Springer Medizin Verlag.
- Reiss, K., Sälzer, C., Schiepe-Tiska, A., Klieme, E. & Köller, O. (Hrsg.). (2016). *PISA 2015. Eine Studie zwischen Kontinuität und Innovation* (1. Auflage, neue Ausgabe). Münster: Waxmann.
- Rekus, J., Hintz, D. & Ladenthin, V. (1998). *Die Hauptschule. Alltag, Reform, Geschichte, Theorie* (Grundlagentexte Pädagogik). Weinheim: Juventa-Verl.
- Reusser, K. & Pauli, C. (2013). Verständnisorientierung in Mathematikstunden erfassen. *Zeitschrift für Pädagogik* 59 (3), 308–335.
- Reyer, T. (2004). *Oberflächenmerkmale und Tiefenstrukturen im Unterricht. Exemplarische Analysen im Physikunterricht der gymnasialen Sekundarstufe* (Studien zum Physiklernen, Bd. 32). Berlin: Logos-Verl.
- Rjosk, C. (2015). *Zuwanderungsbezogene Klassenzusammensetzung. Messung sowie direkte und vermittelte Effekte auf Leistung und psychosoziale Schülermerkmale*

- (Dissertation zur Erlangung des akademischen Grades Doctor rerum naturalium (Dr. rer. nat) im Fach Psychologie).
- Rjosk, C., Richter, D., Hochweber, J., Lüdtke, O., Klieme, E. & Stanat, P. (2014). Socioeconomic and language minority classroom composition and individual reading achievement. The mediating role of instructional quality. *Learning and Instruction* 32, 63–72. doi:10.1016/j.learninstruc.2014.01.007
- Robitzsch, A. (2015). *Essays zu methodischen Herausforderungen im Large-Scale Assessment*.
- Roehler, L. R., Duffy, G. G., Conley, M., Hermann, B. A., Johnson, J. & Michelson, S. (1987). *Exploring preservice teachers' knowledge structures. Paper presented at the annual meeting of the American Educational Research Association (April 1987)*, Washington, DC.
- Roppelt, A., Blum, W. & Pöhlmann, C. (2013). Beschreibung der untersuchten mathematischen Kompetenzen. In H. A. Pant, P. Stanat, U. Schroeders, A. Roppelt, T. Siegle & C. Pöhlmann (Hrsg.), *IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I* (S. 23–37). Münster: Waxmann Verlag.
- Rosenthal, R. & Jacobson, L. (1968). *Pygmalion in the classroom. Teacher expectation and pupils' intellectual development*. New York: Rinehart & Winston.
- Roßmann, J. (2017). *Satisficing in Befragungen. Theorie, Messung und Erklärung* (Schriftenreihe der ASI – Arbeitsgemeinschaft Sozialwissenschaftlicher Institute). Wiesbaden: VS Verlag für Sozialwissenschaften. doi:10.1007/978-3-658-16668-7
- Rothland, M. (2009). Lehrerberuf und Lehrerrolle. In S. Blömeke, T. Bohl, L. Haag, G. Lang-Wojtask & W. Sacher (Hrsg.), *Handbuch Schule. Theorie – Organisation – Entwicklung* (UTB, Bd. 8392). Bad Heilbrunn: Klinkhardt. 494–502.
- Rubie-Davies, C. (2009). Teacher expectations and labeling. In L. J. Saha & A. G. Dworkin (Hrsg.), *International handbook of research on teachers and teaching* (Springer international handbooks of education, Bd. 21, S. 695–707). New York: Springer. doi:10.1007/978-0-387-73317-3_43
- Rumberger, R. W. & Palardy, G. J. (2005). Does segregation still matter? The impact of student composition on academic achievement in high school. *Teachers College Record* 107 (9), 1999–2045.
- Rutkowski, L. & Svetina, D. (2013). Assessing the hypothesis of measurement invariance in the context of large-scale international surveys. *Educational and psychological measurement* 74 (1), 31–57. doi:10.1177/0013164413498257
- Sächsisches Staatsministerium für Kultus. VwV Stundentafeln. <https://www.revosax.sachsen.de/vorschrift/17744>.
- Salisch, M. von. (2016). Die psychologische Perspektive. Persönlichkeitsentwicklung. In S.-M. Köhler, H.-H. Krüger & N. Pfaff (Hrsg.), *Handbuch Peerforschung* (S. 75–87). Opladen: Verlag Barbara Budrich. doi:10.2307/j.ctvd7w8m2.7
- Sälzer, C. & Prenzel, M. (2013). PISA 2012. Eine Einführung in die aktuelle Studie. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (S. 11–45). Münster: Waxmann.
- Sälzer, C., Prenzel, M. & Klieme, E. (2013). Schulische Rahmenbedingungen der Kompetenzentwicklung. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (S. 155–188). Münster: Waxmann.

- Sälzer, C., Reiss, K., Schiepe-Tiska, A., Prenzel, M. & Heinze, A. (2013). Zwischen Grundlagenwissen und Anwendungsbezug. Mathematische Kompetenz im internationalen Vergleich. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (S. 47–97). Münster: Waxmann.
- Scharenberg, K. (2012). *Leistungsheterogenität und Kompetenzentwicklung. Zur Relevanz klassenbezogener Kompositionsmerkmale im Rahmen der KESS-Studie* (Empirische Erziehungswissenschaft, Bd. 36). Münster: Waxmann.
- Scharenberg, K. & Bos, W. (2014). Schulstruktur und Kompetenzentwicklung. Zur Bedeutung von Schulkomposition und Schulformgliederung für die Entwicklung des Leseverständnisses in der Hamburger Sekundarstufe. In H. G. Holtappels (Hrsg.), *Schulentwicklung und Schulwirksamkeit als Forschungsfeld. Theorieansätze und Forschungserkenntnisse zum schulischen Wandel* (S. 141–171). Münster, Westf: Waxmann.
- Schavan, A., Schneider, S., Bos, W., Rösner, E., Klemm, K., Schneider, K. & Dohmen, D. (2008). Ende der Hauptschule. Ausweg aus der Bildungsmisere. *ifo Schnelldienst* 61 (17), 3–20.
- Scheerens, J., Bosker, R. J. & The foundations of educational effectiveness. (1997). *The foundations of educational effectiveness*. Oxford: Pergamon.
- Scheerens, J. (Hrsg.). (2017). *Opportunity to Learn, Curriculum Alignment and Test Preparation. A Research Review* (SpringerBriefs in Education). Cham: Springer International Publishing. doi:10.1007/978-3-319-43110-9
- Schiepe-Tiska, A., Reiss, K., Obersteiner, A., Heine, J.-H., Seidel, T. & Prenzel, M. (2013). Mathematikunterricht in Deutschland. Befunde aus PISA 2012. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (S. 123–154). Münster: Waxmann.
- Schiepe-Tiska, A., Schmidtnr, S., Müller, K., Heine, J.-H., Neumann, K. & Lütke, O. (2016). Naturwissenschaftlicher Unterricht in Deutschland in PISA 2015 im internationalen Vergleich. In K. Reiss, C. Sälzer, A. Schiepe-Tiska, E. Klieme & O. Köller (Hrsg.), *PISA 2015. Eine Studie zwischen Kontinuität und Innovation* (1. Auflage, neue Ausgabe, S. 133–175). Münster: Waxmann.
- Schmidt, H. G. & Boshuizen, H. P. A. (1992). Encapsulation of biomedical knowledge. In D. Evans & V. L. Patel (Hrsg.), *Advanced models of cognitive or medical training and practice* (S. 265–282). New York, NY: Springer. doi:10.1007/978-3-662-02833-9_15
- Schmidt, H. W. & Cogan, L. S. (1996). Development of the TIMSS context questionnaires. In M. O. Martin & D. L. Kelly (Hrsg.), *Third international science study (TIMSS). Technical report, Volume I. Design and development*. Hill, MA: Boston College. <https://timss.bc.edu/timss1995i/TIMSSPDF/TRCHP5.PDF>. Zugegriffen: 16. August 2018.
- Scholl, D. (2011). Der Unterrichtsbegriff in allgemeindidaktischen Theorien. In W. Meseth, M. Proske & F.-O. Radtke (Hrsg.), *Unterrichtstheorien in Forschung und Lehre* (S. 37–49). Bad Heilbrunn: Verlag Julius Klinkhardt.
- Schramm, A. (1975). *Fünfundzwanzig Unterrichtsbegriffe aus dem 20. Jahrhundert. Eine vergleichend-reflektierende Untersuchung zum Problem des Wesens des Unterrichts als Beitrag zur modernen Unterrichtswissenschaft*. München: List Verlag.

- Schubert, V. (2014). Die Vergleichende Erziehungswissenschaft im disziplinären Gefüge der Erziehungswissenschaft. *Enzyklopädie Erziehungswissenschaft Online*, 1–25. doi:10.3262/EEO05140334
- Schuchart, C. & Dunkake, I. (2014). Schichtspezifische Stereotype unter angehenden Lehrkräften. *Zeitschrift für Soziologie der Erziehung und Sozialisation* 34 (1), 89–107.
- Secker, C. E. von & Lissitz, R. W. (1999). Estimating the impact of instructional practices on student achievement in science. *Journal of Research in Science Teaching* 36 (10), 1110–1126. doi:10.1002/(SICI)1098-2736(199912)36:10<1110::AID-TEA4>3.0.CO;2-T
- Seidel, T. (2014). Angebots-Nutzungs-Modelle in der Unterrichtspsychologie. Integration von Struktur- und Prozessparadigma. *Zeitschrift für Pädagogik* 60 (6), 850–866.
- Senkbeil, M., Drechsel, B., Rolff, H.-G., Bensen, M., Zimmer, K., Lehmann, R. H. & Neumann, A. (2004). Merkmale und Wahrnehmungen von Schule und Unterricht. In PISA-Konsortium (Hrsg.), *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland. Ergebnisse des zweiten internationalen Vergleichs* (S. 296–354). Münster: Waxmann.
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher* 29 (7), 4–14. doi:10.3102/0013189X029007004
- Shikano, S. (2010). Einführung in die Inferenz durch den nichtparametrischen Bootstrap. In C. Wolf & H. Best (Hrsg.), *Handbuch der sozialwissenschaftlichen Datenanalyse* (1. Aufl., S. 191–204). Wiesbaden: VS Verlag für Sozialwissenschaften. doi:10.1007/978-3-531-92038-2_9
- Siegler, R., Eisenberg, N., DeLoache J. & Saffran, J. (2016). Beziehung zu Gleichaltrigen. In S. Pauen, R. S. Siegler, N. Eisenberg, J. S. DeLoache & J. Saffran (Hrsg.), *Entwicklungspsychologie im Kindes- und Jugendalter* (4. Auflage, S. 483–528). Berlin: Springer. doi:10.1007/978-3-662-47028-2_13
- Silver, N. C. & Dunlap, W. P. (1987). Averaging correlation coefficients. Should Fisher's z transformation be used? *Journal of Applied Psychology* 72 (1), 146–148. doi:10.1037/0021-9010.72.1.146
- Staatsinstitut für Schulqualität und Bildungsforschung München. (2001). *Realschule R6. genehmigter Lehrplan – gültig für Jgst. 6 bis 10*. <https://www.isb.bayern.de/download/8999/e2.pdf>. Zugriffen: 18. August 2018.
- Staatsinstitut für Schulqualität und Bildungsforschung München. (2010). *Der Lehrplan für das Gymnasium in Bayern im Überblick*. Augsburg: Druckerei Joh. Walch.
- Stäbler, F. (2017). *Die Zusammensetzung der Lerngruppe und ihre Effekte auf psychosoziale Merkmale und Leistung von Schülerinnen und Schülern*. Dissertation. Berlin: Freie Universität Berlin.
- Stäbler, F., Dumont, H., Becker, M. & Baumert, J. (2017). What happens to the fish's achievement in a little pond? A simultaneous analysis of class-average achievement effects on achievement and academic self-concept. *Journal of Educational Psychology* 109 (2), 191–207. doi:10.1037/edu0000135
- Statistisches Bundesamt. (2016). *40 % aller Grundschüler wechseln auf das Gymnasium. Pressemitteilung vom 06. Main 2016*, Wiesbaden. https://www.destatis.de/DE/PresseService/Presse/Pressemitteilungen/2016/05/PD16_154_211pdf.pdf?__blob=publicationFile. Zugriffen: 19. März 2018.
- Statistisches Bundesamt. (2017a). *33% der Schülerinnen und Schüler mit Migrationshintergrund. Zahl der Woche vom 7. Februar 2017*, Wiesbaden. <https://>

- www.destatis.de/DE/PresseService/Presse/Pressemitteilungen/zdw/2017/PD17_006_p002pdf.pdf?__blob=publicationFile. Zugegriffen: 19. März 2018.
- Statistisches Bundesamt (Hrsg.). (2017b). *Bevölkerung und Erwerbstätigkeit. Bevölkerung mit Migrationshintergrund*. Ergebnisse des Mikrozensus 2016 (Fachserie 1 Reihe 2.2), Wiesbaden. https://www.destatis.de/DE/Publikationen/Thematisch/Bevoelkerung/MigrationIntegration/Migrationshintergrund2010220167004.pdf?__blob=publicationFile. Zugegriffen: 28. März 2018.
- Steenkamp, J.-B. E. M. & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research* 25, 78–90. doi:10.1086/209528
- SteG-Konsortium. (2016). *Ganztagsschule. Bildungsqualität und Wirkungen außerunterrichtlicher Angebote*. Ergebnisse der Studie zur Entwicklung von Ganztagsschulen 2012–2015. https://www.projekt-steg.de/sites/default/files/StEG_Brosch_FINA_L.pdf. Zugegriffen: 4. November 2018.
- Steindorf, G. (1981). *Grundbegriffe des Lehrens und Lernens* (5. Aufl.). Bad Heilbrunn: Klinkhardt.
- Steinhoff, A. & Grundmann, M. (2016). Peers und die Reproduktion sozialer Ungleichheit. In S.-M. Köhler, H.-H. Krüger & N. Pfaff (Hrsg.), *Handbuch Peerforschung* (S. 499–513). Opladen: Verlag Barbara Budrich. doi:10.2307/j.ctvd7w8m2.37
- Stichweh, R. (1997). Professions in modern society. *International Review of Sociology* 7 (1), 95–102. doi:10.1080/03906701.1997.9971225
- Stoetzer, M.-W. (2017). *Regressionsanalyse in der empirischen Wirtschafts- und Sozialforschung Band 1. Eine nichtmathematische Einführung mit SPSS und Stata*. Berlin: Springer. doi:10.1007/978-3-662-53824-1
- Swain, M. & Lapkin, S. (2000). Task-based second language learning: the uses of the first language. *Language Teaching Research* 4 (3), 251–274. doi:10.1191/136216800125087
- Tanner, H. (1993). *Einstellungsänderungen während der Lehrerbildung und Berufseinführung. Literaturübersicht und Längsschnittuntersuchung über die berufliche Sozialisation von Schweizer Primarlehrern*. Weinheim: Deutscher Studien Verl.
- Tenorth, H.-E. & Tippelt, R. (2007). *Beltz Lexikon Pädagogik* (Studium Paedagogik). Weinheim: Beltz.
- Terhart, E. (1994). Unterricht. In D. Lenzen (Hrsg.), *Erziehungswissenschaft. Ein Grundkurs* (Rororo Rowohlts Enzyklopädie, S. 133–158). Reinbek: Rowohlt-Taschenbuch-Verlag.
- Terhart, E. (2001). *Lehrerberuf und Lehrerbildung. Forschungsbefunde, Problemanalysen, Reformkonzepte* (Beltz Pädagogik). Weinheim: Beltz Verlagsguppe.
- Terhart, E. (2008). Die Lehrerbildung. In K. S. Cortina, J. Baumert, A. Leschinsky, K. U. Mayer & L. Trommer (Hrsg.), *Das Bildungswesen in der Bundesrepublik Deutschland* (S. 745–772). Reinbek bei Hamburg: Rowohlt-Taschenbuch-Verlag.
- Terhart, E. (2010). Was hat sich in der Lehrerbildung getan? Ein Rückblick. In R. Wernstedt (Hrsg.), *Der Lehrerberuf im Wandel. Wie Reformprozesse Eingang in den Schulalltag finden können* (Schriftenreihe des Netzwerk Bildung, [15], 1. Aufl., S. 11–15). Berlin: Netzwerk Bildung, Friedrich-Ebert-Stiftung.
- Thapa, A., Cohen, J., Guffey, S. & Higgins-D'Alessandro, A. (2013). A review of school climate research. *Review of Educational Research* 83 (3), 375–385. doi:10.3102/0034654313483907

- Thomé, G. & Eichler, W. (2008). Rechtschreiben Deutsch. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Zentrale Befunde der Studie Deutsch Englisch Schülerleistung International (DESI)* (S. 79–104). Weinheim [u.a.]: Beltz.
- Thompson, D. R., Huntley, M. A. & Suurtamm, C. (2017). *International Perspectives on Mathematics Curriculum* (Research in Mathematics Education). Charlotte, NC: Information Age Publishing.
- Thüringer Ministerium für Bildung. (2018). Thüringer Schulordnung. ThürSchulO. <https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=2ahUKEwiX08jFh6TdAhVEgxoKHeNbCj4QFjAAegQIABAC&url=https%3A%2F%2Fwww.thueringen.de%2Fde%2Fpublikationen%2Fpic%2Fpubdownload1245.pdf&usg=AOvVaw1j0-rmaLPzUuB7X60ZaJyA>.
- Tillmann, K.-J. & Meier, U. (2001). Schule, Familie und Freunde – Erfahrungen von Schülerinnen und Schülern in Deutschland. In Deutsches PISA Konsortium (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 468–509). Opladen: Leske + Budrich.
- Tillmann, K.-J., Dederig, K., Kneuper, D., Kuhlmann, C. & Nessel, I. (2008). *PISA als bildungspolitisches Ereignis. Fallstudien in vier Bundesländern* (Schule und Gesellschaft, Bd. 43, 1. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften / GWV Fachverlage. doi:10.1007/978-3-531-91127-4
- Todorov, T. & Böhringer, W. (2008). *Die Eroberung Amerikas. Das Problem des Anderen* (Edition Suhrkamp, 1213 = N.F., 213, [Nachdr.]. Frankfurt am Main: Suhrkamp.
- Trautsch, M. (2016, 24. September). Die Hauptschule als Chance für Deutschkurse für Flüchtlinge. *Frankfurter Allgemeine*. <https://www.faz.net/aktuell/rhein-main/die-hauptschule-als-chance-fuer-fluechtlinge-14205555.html>. Zugriffen: 7. Dezember 2018.
- Trautwein, U. & Neumann, M. (2008). Das Gymnasium. In K. S. Cortina, J. Baumert, A. Leschinsky, K. U. Mayer & L. Trommer (Hrsg.), *Das Bildungswesen in der Bundesrepublik Deutschland* (S. 467–501). Reinbek bei Hamburg: Rowohlt-Taschenbuch-Verlag. doi:10.30965/9783657764969_018
- Tucker, L. R. (1951). *A method for synthesis of factor analysis studies. Personnel Research Section Report No. 984*. Washington, DC: Department of the Army.
- Turner, J. C., Christensen, A. & Meyer, D. K. (2009). Teachers' beliefs about student learning and motivation. In L. J. Saha & A. G. Dworkin (Hrsg.), *International handbook of research on teachers and teaching* (Springer international handbooks of education, Bd. 21, S. 361–371). Boston, MA: Springer US. doi:10.1007/978-0-387-73317-3_23
- Uchida, Y., Norasakkunkit, V. & Kitayama, S. (2004). Cultural constructions of happiness. Theory and empirical evidence. *Journal of Happiness Studies* 5 (3), 223–239. doi:10.1007/s10902-004-8785-9
- Unesco. (1997). *International Standard Classification of Education. ISCED 1997*. http://www.unesco.org/education/information/nfsunesco/doc/isced_1997.htm. Zugriffen: 2. November 2018.
- Urban, D. & Mayerl, J. (2018). *Angewandte Regressionsanalyse: Theorie, Technik und Praxis* (Studienskripten zur Soziologie, 5., überarbeitete Auflage). Wiesbaden: Springer VS. doi:10.1007/978-3-658-01915-0

- van Ackeren, I. & Klemm, K. (2011). *Entstehung, Struktur und Steuerung des deutschen Schulsystems. Eine Einführung* (2., aktualisierte und überarbeitete Auflage). Wiesbaden: VS Verlag für Sozialwissenschaften. doi:10.1007/978-3-531-92813-5
- van de Vijver, F. J. R. (1998). Towards a theory of bias and equivalence. In J. A. Harkness (Hrsg.), *Cross-cultural survey equivalence* (Nachrichten Spezial, Bd. 3, S. 41–65). Mannheim: Zentrum für Umfragen, Methoden und Analysen (ZUMA).
- van de Vijver, F. J. R. & Leung, K. (1998). *Methods and data analysis for cross-cultural research* (Cross-cultural psychology series, Bd. 1, 2. Dr.). Thousand Oaks, Calif. Sage Publications.
- van de Vijver, F. J. R. & Leung, K. (2011). Equivalence and bias. A review of concepts, models, and data analytic procedures. In D. R. Matsumoto (Hrsg.), *Cross-cultural research methods in psychology* (Culture and psychology, S. 17–45). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511779381.003
- van de Vijver, F. J. R. & Poortinga, Y. H. (1997). Towards an integrated analysis of bias in cross-cultural assessment. *European Journal of Psychological Assessment* 13 (1), 29–37. doi:10.1027/1015-5759.13.1.29
- van Ewijk, R. & Sleegers, P. (2010). The effect of peer socioeconomic status on student achievement. A meta-analysis. *Educational Research Review* 5 (2), 134–150. doi:10.1016/j.edurev.2010.02.001
- van Herk, H., Poortinga, Y. H. & Verhallen, T. M. M. (2016). Response Styles in Rating Scales. *Journal of Cross-Cultural Psychology* 35 (3), 346–360. doi:10.1177/0022022104264126
- van Ophuysen, S. & Wendt, H. (2010). Zur Veränderung der Mathematikleistung von Klasse 4 bis 6. Welchen Einfluss haben Kompositions- und Unterrichtsmerkmale. In J. Baumert, K. Maaz & U. Trautwein (Hrsg.), *Bildungsentscheidungen. Zeitschrift für Erziehungswissenschaft Sonderheft 12/2009* (S. 302–327). Wiesbaden: VS Verlag für Sozialwissenschaften. doi:10.1007/978-3-531-92216-4_13
- van Someren, M. W., Barnard, Y. F. & Sandberg, J. A. C. (1994). *The think aloud method. A practical guide to modelling cognitive processes*. London: Academic Press.
- Vandenberghe, R. & Huberman, A. M. (1999). *Understanding and preventing teacher burnout. A sourcebook of international research and practice* (The Jacobs Foundation series on adolescence). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511527784
- Vieluf, U., Ivanov, S. & Nikolova, R. (2011). *Kompetenzen und Einstellungen von Schülerinnen und Schülern an Hamburger Schulen am Ende der Sekundarstufe I und zu Beginn der gymnasialen Oberstufe. KESS 10/11* (Freie und Hansestadt Hamburg – Behörde für Schule und Berufsbildung, Hrsg.), Hamburg.
- Vieluf, S., Hochweber, J., Klieme, E. & Kunter, M. (2015). Who has a good relationship with the teachers? A comparison of comprehensive education systems with education systems using between-school tracking. *Oxford Review of Education* 41 (1), 3–25. doi:10.1080/03054985.2014.992874
- Völk, K. (2014). *Die Rolle der Realschule im gegliederten Schulsystem* (Berichte aus der Pädagogik). Aachen: Shaker Verlag (Dissertation).
- Voss, T., Kleickmann, T., Kunter, M. & Hachfeld, A. (2011). Überzeugungen von Mathematiklehrkräften. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 235–257). Münster: Waxmann.

- Wagner, R. F. (2016). *Unterricht aus Sicht der Lehrerinnen und Lehrer. Subjektive Theorien zur Unterrichtsgestaltung und ihre Veränderung durch ein Training zu neuen Unterrichtsmethoden* (Klinkhardt forschung). Bad Heilbrunn: Julius Klinkhardt.
- Wagner, W., Helmke, A., Schrader, F.-W., Eichler, W., Thomé, G. & Willenberg, H. (2008). Selbstkonzept und Motivation im Fach Deutsch. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Zentrale Befunde der Studie Deutsch Englisch Schülerleistung International (DESI)* (S. 231–243). Weinheim [u.a.]: Beltz.
- Wahrig, G. & Wahrig-Burfeind, R. (2003). *Wahrig Fremdwörterlexikon*. Gütersloh: Bertelsmann Lexikon Institut.
- Weijters, B. (2006). *Response styles in consumer research* (Dissertation submitted to the Faculty of Economics and Business Administration, Ghent University).
- Weinert, W. (1997). In M. A. Meyer (Hrsg.), *Lernmethoden. Lehrmethoden. Wege zur Selbständigkeit* (Friedrich-Jahresheft, 15, S. 50–52). Seelze: Erhard Friedrich.
- Weippert, G. (1964). *Jenseits von Individualismus und Kollektivismus. Studien zum gegenwärtigen Zeitalter*. Düsseldorf: Schilling.
- Wendt, H., Gröhllich, C., Guill, K., Scharenberg, K. & Bos, W. (2010). Die Kompetenzen der Schülerinnen und Schüler im Leseverständnis. In W. Bos & C. Gröhllich (Hrsg.), *KESS 8. Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 8* (HANSE – Hamburger Schriften zur Qualität im Bildungswesen, Bd. 6, S. 21–36). Münster: Waxmann.
- Wiater, W. (2014). *Unterrichtsprinzipien. Prüfungswissen – Basiswissen Schulpädagogik* (Immer besser unterrichten Didaktik, 6., überarbeitete Auflage). Donauwörth: Auer Verlag.
- Wiechmann, J. & Wildhirt, S. (2016). Unterrichtsmethoden. vom Nutzen der Vielfalt. In J. Wiechmann & S. Wildhirt (Hrsg.), *Zwölf Unterrichtsmethoden. Vielfalt für die Praxis* (Beltz Pädagogik, 6., vollständig überarbeitete Auflage, S. 11–23). Weinheim: Beltz.
- Wilhelm, O., Schroeders, U. & Schipolowski, S. (2014). *Berliner Test zur Erfassung fluider und kristaliner Intelligenz für die 8. bis 10. Jahrgangsstufe*. Göttingen: Hogrefe.
- Wiliam, D. & Thompson, M. (2008). Integrating assessment with learning. What will it take to make it work? In C. A. Dwyer (Hrsg.), *The future of assessment. Shaping teaching and learning* (S. 53–82). New York: Lawrence Erlbaum. doi:10.4324/9781315086545-3
- Wilkinson, I. A. (2002). Introduction. Peer influences on learning: where are they? *International Journal of Educational Research* 37 (5), 395–401. doi:10.1016/S0883-0355(03)00012-0
- Willenberg, H. (2008). Wortschatz Deutsch. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Zentrale Befunde der Studie Deutsch Englisch Schülerleistung International (DESI)* (S. 72–80). Weinheim [u.a.]: Beltz.
- Willis, P. (1982). *Spaß am Widerstand. Gegenkultur in der Arbeiterschule* (2. Aufl.). Frankfurt am Main: Syndikat.
- Wolf, C. & Best, H. (Hrsg.). (2010). *Handbuch der sozialwissenschaftlichen Datenanalyse* (1. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften. doi:10.1007/978-3-531-92038-2

- Wolf, E. J., Harrington, K. M., Clark, S. L. & Miller, M. W. (2013). Sample Size Requirements for Structural Equation Models: An Evaluation of Power, Bias, and Solution Propriety. *Educational and psychological measurement* 76 (6), 913–934. doi:10.1177/0013164413495237
- Wollenweber, H. (1997). *Die Realschule in Geschichte und Gegenwart* (Studien und Dokumentationen zur deutschen Bildungsgeschichte, Bd. 66). Köln [u.a.]: Böhlau.
- Youniss, J. (1982). Die Entwicklung und Funktion von Freundschaftsbeziehungen. In W. Edelstein & M. Keller (Hrsg.), *Perspektivität und Interpretation. Beiträge zur Entwicklung des sozialen Verstehens* (Suhrkamp Taschenbuch Wissenschaft, Bd. 364, S. 78–108). Frankfurt am Main: Suhrkamp.
- Zurbriggen, C. (2016). *Schulklasseneffekte*. Wiesbaden: Springer Fachmedien Wiesbaden.
- Zymek, B. (2013). Die Zukunft des zweigliedrigen Schulsystems in Deutschland. Was man von der historischen Schulentwicklung dazu wissen kann. *Zeitschrift für Pädagogik* 59 (4), 469–481.

10 Abbildungsverzeichnis

Abbildung 1. Gegenüberstellung der Schulartbeschreibungen.....	28
Abbildung 2. Die Schularten der Sekundarstufe I.....	43
Abbildung 3. Didaktisches Dreieck.....	45
Abbildung 4. Didaktisches Dreieck mit schulartspezifischen Ausgangsbedingungen.....	46
Abbildung 5. Das Kompetenzmodell von COACTIV.....	59
Abbildung 6. Gegenüberstellung konfiguraler Invarianz und konfiguraler Varianz	143
Abbildung 7. Vereinfachtes Mediationsmodell der interessierenden Effekte auf Klassenebene	155
Abbildung 8. Vollständiges Mediationsmodell mit Kontrolle von Antworttendenzen	157
Abbildung 9. Vereinfachtes Modell eines Interaktionseffekts von Schulart und Unterrichtsprozessmerkmal auf Klassenebene	159
Abbildung 10. Vollständiges Mehrebenen-Moderationsmodell mit Kontrolle der Leistung auf Individualebene und Kontrolle von Antworttendenzen.....	161
Abbildung 11. Mittlere Zusammensetzung der Klassen nach kognitiven Merkmalen in den Schularten.....	189
Abbildung 12. Mittlere Zusammensetzung der Klassen nach sozioökonomischen Merkmalen in den Schularten	194
Abbildung 13. Mittlere Zusammensetzung der Klassen nach affektiven Merkmalen in den Schularten.....	196
Abbildung 14. Überzeugungen der Lehrenden zu den Zielen des Mathematikunterrichts in den Schularten.....	199
Abbildung 15. Inhalte des Mathematikunterrichts in den Schularten.....	202
Abbildung 16. Unterrichtsprozesse im Fach Mathematik in den Schularten.....	205

Abbildung 17. Zusammenfassende Darstellung bedeutender Schulartunterschiede.....	215
Abbildung 18. Mediation der Schulartunterschiede im Potential zur kognitiven Aktivierung durch Variation der Aufgaben über mathematische Konzepte.....	219
Abbildung 19. Mediation der Schulartunterschiede in schülerorientierten Unterrichtsmethoden über mathematische Konzepte	220
Abbildung 20. Mediation der Schulartunterschiede in schülerorientierten Unterrichtsmethoden über reine Mathematik	221
Abbildung 21. Mediation der Schulartunterschiede in schülerorientierten Unterrichtsmethoden über die mittlere Klassenleistung im PISA-Mathematiktest	222
Abbildung 22. Mediationsmodell der Schulartunterschiede in schülerorientierten Unterrichtsmethoden über die Merkmale mathematische Konzepte, Leistungskomposition und reine Mathematik.....	223
Abbildung 23. Zusammenfassung der Schulartunterschiede.....	235
Abbildung 24. Zuordnung der Organisationformen allgemeiner Schulen der Sekundarstufe I zu den Schularten in PISA 2012 (Online-Anhang).....	297
Abbildung 25. Lehrerüberzeugungen zu den Zielen des Mathematikunterrichts in den Schularten (Online-Anhang)...	315
Abbildung 26. Mittlere Zusammensetzung der Klassen nach affektiven Merkmale in den Schularten (Online-Anhang)	317
Abbildung 27. Unterrichtsprozesse im Fach Mathematik in den Schularten (Online-Anhang).....	319

11 Tabellenverzeichnis

Tabelle 1. Verteilung der Lernenden in 9. Klassen auf die Schularten in Prozent.....	36
Tabelle 2. Schulabgänger aus integrierten Gesamtschulen in Deutschland nach Abschlussart im Schuljahr 2015/2016.....	39
Tabelle 3. Schulabgänger aus Schulen mit zwei Bildungsgängen in Deutschland nach Abschlussart im Schuljahr 2015/2016.....	42
Tabelle 4. Skalen, Indizes und Leistungsdaten der Lernenden aus PISA 2012.....	136
Tabelle 5. Skalen der Lehrenden aus PISA 2012 – nationale Erweiterung	139
Tabelle 6. Ergebnisse der Invarianzanalyse.....	166
Tabelle 7. Faktorstruktur der Ziele im Mathematikunterricht für die Gesamtstichprobe	170
Tabelle 8. Finale Dimensionen der Ziele im Mathematikunterricht für die Gesamtstichprobe	173
Tabelle 9. Kontingenzkoeffizienten für die Dimensionen der Ziele im Mathematikunterricht	174
Tabelle 10. Faktorstruktur des Potentials zur kognitiven Aktivierung für die Gesamtstichprobe	176
Tabelle 11. Finale Dimensionen des Potentials zur kognitiven Aktivierung für die Gesamtstichprobe.....	177
Tabelle 12. Kontingenzkoeffizienten für die Dimensionen des Potentials zur kognitiven Aktivierung.....	178
Tabelle 13. Tendenz zur Zustimmung und zur Wahl von mittleren Kategorien pro Schulart und gesamt	179
Tabelle 14. Veränderung der Werte bei Überzeugungsmerkmalen der Lehrenden unter Kontrolle von Antworttendenzen	182
Tabelle 15. Veränderung der Werte bei affektiven Schülermerkmalen unter Kontrolle von Antworttendenzen	183
Tabelle 16. Veränderung der Werte bei Unterrichtsmerkmalen unter Kontrolle von Antworttendenzen	183

Tabelle 17. Analysierte Unterrichtsprozessmerkmale und potentielle Mediatoren auf Seiten der Lehrenden, der Unterrichtsinhalte und der Schülergruppe.....	217
Tabelle 18. Moderierende Effekte der Schulart auf den Zusammenhang von Unterrichtsprozessmerkmalen und der Leistungsentwicklung gemessen mit dem BiSta-Test	228
Tabelle 19. Moderierende Effekte der Schulart auf den Zusammenhang von Unterrichtsprozessmerkmalen und der Leistungsentwicklung gemessen mit dem PISA-Test	230
Tabelle 20. Schularten der Sekundarstufe I in den Ländern außer Gymnasium (Online-Anhang)	296
Tabelle 21. Lernende in Deutschland in Schulen mit mehreren Bildungsgängen im Schuljahr 2015/2016 (Online-Anhang)	298
Tabelle 22. Faktorstruktur der Ziele im Mathematikunterricht für Schulen mit mehreren Bildungsgängen (Online-Anhang)	299
Tabelle 23. Finale Dimensionen der Ziele im Mathematikunterricht für Schulen mit mehreren Bildungsgängen (Online-Anhang).....	300
Tabelle 24. Faktorstruktur der Ziele im Mathematikunterricht für Realschulen (Online-Anhang)	301
Tabelle 25. Finale Dimensionen der Ziele im Mathematikunterricht für Realschulen (Online-Anhang)	302
Tabelle 26. Faktorstruktur der Ziele im Mathematikunterricht für integrierte Gesamtschulen (Online-Anhang).....	303
Tabelle 27. Finale Dimensionen der Ziele im Mathematikunterricht für integrierte Gesamtschulen (Online-Anhang).....	304
Tabelle 28. Faktorstruktur der Ziele im Mathematikunterricht für Gymnasien (Online-Anhang)	305
Tabelle 29. Finale Dimensionen der Ziele im Mathematikunterricht für Gymnasien (Online-Anhang)	306
Tabelle 30. Faktorstruktur des Potentials zu kognitiven Aktivierung für Schulen mit mehreren Bildungsgängen (Online-Anhang).....	307

Tabelle 31. Finale Dimensionen des Potentials zur kognitiven Aktivierung für Schulen mit mehreren Bildungsgängen (Online-Anhang).....	308
Tabelle 32. Faktorstruktur des Potentials zu kognitiven Aktivierung für Realschulen (Online-Anhang)	309
Tabelle 33. Finale Dimensionen des Potentials zur kognitiven Aktivierung für Realschulen (Online-Anhang)	310
Tabelle 34. Faktorstruktur des Potentials zu kognitiven Aktivierung für integrierte Gesamtschulen (Online-Anhang).....	311
Tabelle 35. Finale Dimensionen des Potentials zur kognitiven Aktivierung für integrierte Gesamtschulen (Online-Anhang)	312
Tabelle 36. Faktorstruktur des Potentials zu kognitiven Aktivierung für Gymnasien (Online-Anhang)	313
Tabelle 37. Finale Dimensionen des Potentials zur kognitiven Aktivierung für Gymnasien (Online-Anhang).....	314
Tabelle 38. Mediation der Schulartunterschiede im formativen Assessment (Online-Anhang).....	321
Tabelle 39. Mediation der Schulartunterschiede in schülerorientierten Unterrichtsmethoden (Online-Anhang)	322
Tabelle 40. Mediation der Schulartunterschiede im Potential zur kognitiven Aktivierung durch Insistieren auf Begründen (Online-Anhang)	323
Tabelle 41. Mediation der Schulartunterschiede im Potential zur kognitiven Aktivierung durch Variation der Aufgaben (Online-Anhang)	324

ANHANG

Inhalt

12	Anhang	296
12.1	Schularten der Sekundarstufe I	296
12.2	Zuordnung der Organisationsformen zu den Schularten.....	297
12.3	Lernende in Schulen mit mehreren Bildungsgängen im Schuljahr 2015/2016.....	298
12.4	Faktorstrukturen innerhalb der Schularten	299
12.4.1	Ziele im Mathematikunterricht.....	299
12.4.1.1	Schulen mit mehreren Bildungsgängen	299
12.4.1.2	Realschulen	301
12.4.1.3	Integrierte Gesamtschulen	303
12.4.1.4	Gymnasien	305
12.4.2	Potential zur kognitive Aktivierung	307
12.4.2.1	Schulen mit mehreren Bildungsgängen	307
12.4.2.2	Realschulen	309
12.4.2.3	Integrierte Gesamtschulen	311
12.4.2.4	Gymnasien	313
12.5	Auswirkung von Antworttendenzen.....	315
12.5.1	Lehrerüberzeugungen	315
12.5.2	Affektive Merkmale	317
12.5.3	Unterrichtsprozesse	319
12.6	Mediationsmodelle	321
12.6.1	Mediation der Schulartunterschiede im formativen Assessment	321
12.6.2	Mediation der Schulartunterschiede in schülerorientierten Unterrichtsmethoden	322
12.6.3	Mediation der Schulartunterschiede im Potential zur kognitiven Aktivierung durch Insistieren auf Begründen	323
12.6.4	Mediation der Schulartunterschiede im Potential zur kognitiven Aktivierung durch Variation der Aufgaben.....	324

12 Anhang

12.1 Schularten der Sekundarstufe I¹

Tabelle 20. Schularten der Sekundarstufe I in den Ländern außer Gymnasium

SEKRETARIAT DER STÄNDIGEN KONFERENZ
DER KULTUSMINISTER DER LÄNDER
IN DER BUNDESREPUBLIK DEUTSCHLAND

Berlin, den 08.02.2018
Tel. (030) 25418-439
schulen@kmk.org
www.kmk.org

- II A/Sw -

Schularten der Sekundarstufe I in den Ländern außer Gymnasium¹ (Stand: Februar 2018)

Land	Hauptschule	Realschule	Schularten mit zwei Bildungsgängen	Schularten mit drei Bildungsgängen
BW	X + Werkrealschule	X		Gesamtschule an drei Standorten Gemeinschaftsschule
BY	X Mittelschule	X		
BE	entfällt ab 01.08.2011	entfällt ab 01.08.2011		Integrierte Sekundarschule
BB			Oberschule	Gesamtschule
HB				Oberschule
HH				Stadtteilschule
HE	X	X	Verbundene Haupt- und Realschule (ab dem 01.08.2011 Mittelstufenschule)	Gesamtschule
MV			Regionale Schule	Gesamtschule Regionale Schule teilweise
NI	X	X		Gesamtschule Oberschule (kann auch nur zwei Bildungsgänge umfassen)
NW	X	X		Gesamtschule Sekundarschule (seit Oktober 2011)
RP	X (bis 31.07.2013)	X (bis 31.07.2013)	Realschule plus	Gesamtschule
SL			Erweiterte Realschule (läuft ab 2012/13 aus)	Gesamtschule (läuft ab 2012/13 aus) Gemeinschaftsschule (ab 2012/13 im Aufbau)
SN			Mittelschule (seit 01.08.2013 unter der Bezeichnung Oberschule)	
ST			Sekundarschule	Gesamtschule Gemeinschaftsschule
SH	entfällt ab 01.08.2011	entfällt ab 01.08.2011	Regionalschule	Gemeinschaftsschule
TH			Regelschule	Gesamtschule Gemeinschaftsschule

¹ Das Gymnasium gibt es in jedem Land.

1 Unveröffentlichtes Schreiben des Sekretariats der Ständigen Konferenz der Kultusministerien der Länder in der Bundesrepublik Deutschland vom 08. Februar 2018.

12.2 Zuordnung der Organisationsformen zu den Schularten

Schulart	Organisationsform
Hauptschule	Hauptschule/ Volksschule
Realschule	Realschule (und Wirtschaftsschule in BY)
Gymnasium	Gymnasium
Integrierte Gesamtschule	Integrative Gesamtschule/ Gemeinschaftsschule
	Waldorfschule
Schule mit mehreren Bildungsgängen	Kooperative Gesamtschule <i>Hauptschulzweig</i>
	Kooperative Gesamtschule <i>Realschulzweig</i>
	Kooperative Gesamtschule <i>gymnasialer Zweig</i>
	Regelschule/ Mittelschule/ Regionalschule/ Regionale Schule/ Mittelstufenschule/ Sekundarschule/ Erweiterte Realschule/ Integrierte Haupt- und Realschule/ (Duale) Oberschule/ Realschule plus/ Integrierte Sekundarschule/ Stadtteilschule/ Werkrealschule: <i>voraussichtlicher Hauptschulabschluss</i>
	Regelschule/ Mittelschule/ Regionalschule/ Regionale Schule/ Mittelstufenschule/ Sekundarschule/ Erweiterte Realschule/ Integrierte Haupt- und Realschule/ (Duale) Oberschule/ Realschule plus/ Integrierte Sekundarschule/ Stadtteilschule/ Werkrealschule: <i>voraussichtlicher Realschulabschluss</i>
	Regelschule/ Mittelschule/ Regionalschule/ Regionale Schule/ Mittelstufenschule/ Sekundarschule/ Erweiterte Realschule/ Integrierte Haupt- und Realschule/ (Duale) Oberschule/ Realschule plus/ Integrierte Sekundarschule/ Stadtteilschule/ Werkrealschule: <i>Hauptschulklass</i>
	Regelschule/ Mittelschule/ Regionalschule/ Regionale Schule/ Mittelstufenschule/ Sekundarschule/ Erweiterte Realschule/ Integrierte Haupt- und Realschule/ (Duale) Oberschule/ Realschule plus/ Integrierte Sekundarschule/ Stadtteilschule/ Werkrealschule: <i>Realschulklass</i>

Abbildung 24. Zuordnung der Organisationformen allgemeiner Schulen der Sekundarstufe I zu den Schularten in PISA 2012.

Die Tabelle wurde auf Basis interner Dokumente erstellt: der Study Programm Table (SPT) für PISA 2012 sowie der Syntax „Step_002_PISA_plus - gradebased_sample_Zusaetzliche Variablen erstellen“ des nationalen PISA-Plus-Datensatzes.

12.3 Lernende in Schulen mit mehreren Bildungsgängen im Schuljahr 2015/2016

Tabelle 21. Lernende in Deutschland in Schulen mit mehreren Bildungsgängen im Schuljahr 2015/2016.

Schüler in MBG 2015/2016		
Bundesland	N	%
Baden-Württemberg	---	---
Bayern	---	---
Berlin	---	---
Brandenburg	32397	6,56
Bremen	814	0,16
Hamburg	---	---
Hessen	550	0,11
Mecklenburg-Vorpommern	42139	8,53
Niedersachsen	74564	15,10
Nordrhein-Westfalen	43610	8,83
Rheinland-Pfalz	85326	17,28
Saarland	5248	1,06
Sachsen	103762	21,01
Sachsen-Anhalt	46033	9,32
Schleswig-Holstein	13999	2,83
Thüringen	45394	9,19
Gesamt	493836	100,00

Quelle: Statistisches Bundesamt GENESIS-Online Datenbank (2018): Schüler: Bundesländer, Schuljahr, Geschlecht, Schulart. Eigene Berechnung der Prozente.

12.4 Faktorstrukturen innerhalb der Schularten

12.4.1 Ziele im Mathematikunterricht

12.4.1.1 Schulen mit mehreren Bildungsgängen

Tabelle 22. Faktorstruktur der Ziele im Mathematikunterricht für Schulen mit mehreren Bildungsgängen.

Item	Beherrschen von Routinen	Modellierungsfähigkeit	Anwendung im Alltag	Interesse
1 Sicherheit bei der Bearbeitung von Standardaufgaben	-0,07	0,81	0,04	0,25
2 Vermittlung von mathematischem Grundwissen	0,02	0,72	0,31	-0,20
3 Sichere Beherrschung von mathematischen Routineoperationen	-0,16	0,58	0,27	0,41
4 Automatisierte Beherrschung der wichtigsten Algorithmen der Mittelstufe	0,29	0,72	-0,21	0,16
5 Schnelles Rechnen im Kopf bei Standardanwendungen	-0,08	0,15	-0,08	0,82
6 Mathematisierungsfähigkeit entwickeln	0,58	-0,25	0,32	0,06
7 Alltagssituationen in mathematische Modelle zu übersetzen	0,54	0,12	0,36	-0,14
8 Mathematisch argumentieren können	0,70	-0,20	0,09	0,09
9 Die Angemessenheit eines mathematischen Modells erkennen können	0,70	-0,25	-0,02	0,25
10 Interesse für die Anwendungen der Mathematik	0,27	0,03	0,14	0,53
11 Mathematik im Alltag verstehen	0,06	0,07	0,86	0,02
12 Mathematisches Wissen im Alltag anwenden können	-0,03	0,03	0,84	0,02
13 Freude am mathematischen Denken entwickeln	0,83	0,31	-0,07	-0,12
14 Interesse für die Anwendungen der Mathematik	0,85	0,24	-0,08	-0,06
15 Mathematik zu etwas persönlich Wichtigem machen	0,34	0,05	0,30	0,47
16 Erreichen, dass die Schülerinnen und Schüler Spaß an der Mathematik haben	0,50	0,00	0,07	0,34

Tabelle 23. Finale Dimensionen der Ziele im Mathematikunterricht für Schulen mit mehreren Bildungsgängen.

Item	Beherrschen von Routinen	Modellierungsfähigkeit	Anwendung im Alltag
1 Sicherheit bei der Bearbeitung von Standardaufgaben	0,91	-0,06	0,02
2 Vermittlung von mathematischem Grundwissen	0,58	-0,21	-0,38
3 Sichere Beherrschung von mathematischen Routineoperationen	0,74	-0,04	-0,18
4 Automatisierte Beherrschung der wichtigsten Algorithmen der Mittelstufe	0,85	0,23	0,26
6 Mathematisierungsfähigkeit entwickeln	-0,10	0,73	-0,21
7 Alltagssituationen in mathematische Modelle zu übersetzen	0,19	0,62	-0,24
8 Mathematisch argumentieren können	0,00	0,84	0,01
9 Die Angemessenheit eines mathematischen Modells erkennen können	0,00	0,85	0,10
11 Mathematik im Alltag verstehen	0,05	0,21	-0,82
12 Mathematisches Wissen im Alltag anwenden können	-0,02	0,08	-0,85

12.4.1.2 Realschulen

Tabelle 24. Faktorstruktur der Ziele im Mathematikunterricht für Realschulen.

Item	Beherrschen von Routinen	Modellierungsfähigkeit	Anwendung im Alltag	Interesse
1 Sicherheit bei der Bearbeitung von Standardaufgaben	-0,02	0,80	0,06	-0,13
2 Vermittlung von mathematischem Grundwissen	-0,11	0,75	0,22	0,06
3 Sichere Beherrschung von mathematischen Routineoperationen	-0,07	0,89	-0,07	0,05
4 Automatisierte Beherrschung der wichtigsten Algorithmen der Mittelstufe	0,40	0,62	-0,27	0,06
5 Schnelles Rechnen im Kopf bei Standardanwendungen	0,69	0,11	-0,26	0,08
6 Mathematisierungsfähigkeit entwickeln	0,68	-0,19	-0,01	0,21
7 Alltagssituationen in mathematische Modelle zu übersetzen	0,80	-0,07	0,40	-0,18
8 Mathematisch argumentieren können	0,74	0,04	0,18	0,04
9 Die Angemessenheit eines mathematischen Modells erkennen können	0,67	0,01	0,14	0,14
10 Interesse für die Anwendungen der Mathematik	0,03	-0,01	0,18	0,82
11 Mathematik im Alltag verstehen	0,14	0,06	0,65	0,42
12 Mathematisches Wissen im Alltag anwenden können	0,26	0,11	0,81	0,05
13 Freude am mathematischen Denken entwickeln	0,19	0,01	-0,09	0,73
14 Interesse für die Anwendungen der Mathematik	-0,09	0,01	0,26	0,62
15 Mathematik zu etwas persönlich Wichtigem machen	0,44	-0,07	-0,15	0,50
16 Erreichen, dass die Schülerinnen und Schüler Spaß an der Mathematik haben	0,00	0,02	-0,15	0,73

Tabelle 25. Finale Dimensionen der Ziele im Mathematikunterricht für Realschulen.

Item	Beherrschen von Routinen	Modellierungsfähigkeit	Anwendung im Alltag
1 Sicherheit bei der Bearbeitung von Standardaufgaben	0,79	-0,14	-0,10
2 Vermittlung von mathematischem Grundwissen	0,75	-0,08	-0,20
3 Sichere Beherrschung von mathematischen Routineoperationen	0,89	-0,03	0,06
4 Automatisierte Beherrschung der wichtigsten Algorithmen der Mittelstufe	0,62	0,51	0,33
6 Mathematisierungsfähigkeit entwickeln	-0,18	0,80	0,00
7 Alltagssituationen in mathematische Modelle zu übersetzen	-0,08	0,68	-0,32
8 Mathematisch argumentieren können	0,04	0,80	-0,12
9 Die Angemessenheit eines mathematischen Modells erkennen können	0,02	0,78	-0,07
11 Mathematik im Alltag verstehen	0,07	0,20	-0,80
12 Mathematisches Wissen im Alltag anwenden können	0,11	0,13	-0,88

12.4.1.3 Integrierte Gesamtschulen

Tabelle 26. Faktorstruktur der Ziele im Mathematikunterricht für integrierte Gesamtschulen.

Item	Beherrschen von Routinen	Modellierungsfähigkeit	Anwendung im Alltag	Interesse
1 Sicherheit bei der Bearbeitung von Standardaufgaben	0,91	-0,19	0,13	0,04
2 Vermittlung von mathematischem Grundwissen	0,70	0,09	-0,24	-0,01
3 Sichere Beherrschung von mathematischen Routineoperationen	0,89	0,06	-0,11	-0,03
4 Automatisierte Beherrschung der wichtigsten Algorithmen der Mittelstufe	0,54	0,00	0,32	0,20
5 Schnelles Rechnen im Kopf bei Standardanwendungen	0,24	0,62	0,05	0,01
6 Mathematisierungsfähigkeit entwickeln	0,20	-0,22	0,73	-0,14
7 Alltagssituationen in mathematische Modelle zu übersetzen	0,20	-0,07	0,50	-0,13
8 Mathematisch argumentieren können	-0,22	0,01	0,74	0,23
9 Die Angemessenheit eines mathematischen Modells erkennen können	-0,10	0,26	0,74	0,12
10 Interesse für die Anwendungen der Mathematik	0,32	0,08	-0,08	0,62
11 Mathematik im Alltag verstehen	-0,13	0,91	-0,14	0,01
12 Mathematisches Wissen im Alltag anwenden können	-0,19	0,93	-0,01	0,05
13 Freude am mathematischen Denken entwickeln	0,15	0,00	-0,56	0,60
14 Interesse für die Anwendungen der Mathematik	-0,10	-0,16	0,16	0,53
15 Mathematik zu etwas persönlich Wichtigem machen	-0,07	0,18	0,12	0,64
16 Erreichen, dass die Schülerinnen und Schüler Spaß an der Mathematik haben	0,09	0,00	-0,11	0,72

Tabelle 27. Finale Dimensionen der Ziele im Mathematikunterricht für integrierte Gesamtschulen.

Item	Beherrschen von Routinen	Modellierungsfähigkeit	Anwendung im Alltag
1 Sicherheit bei der Bearbeitung von Standardaufgaben	0,88	-0,13	0,21
2 Vermittlung von mathematischem Grundwissen	0,76	-0,25	-0,09
3 Sichere Beherrschung von mathematischen Routineoperationen	0,90	-0,12	-0,07
4 Automatisierte Beherrschung der wichtigsten Algorithmen der Mittelstufe	0,56	0,40	-0,01
6 Mathematisierungsfähigkeit entwickeln	0,12	0,73	-0,19
7 Alltagssituationen in mathematische Modelle zu übersetzen	0,20	0,69	-0,62
8 Mathematisch argumentieren können	-0,22	0,80	-0,01
9 Die Angemessenheit eines mathematischen Modells erkennen können	-0,08	0,74	0,34
11 Mathematik im Alltag verstehen	-0,01	0,15	-0,94
12 Mathematisches Wissen im Alltag anwenden können	-0,07	-0,01	-0,94

12.4.1.4 Gymnasien

Tabelle 28. Faktorstruktur der Ziele im Mathematikunterricht für Gymnasien.

Item	Beherrschen von Routinen	Modellierungsfähigkeit	Anwendung im Alltag	Interesse
1 Sicherheit bei der Bearbeitung von Standardaufgaben	-0,07	0,82	0,03	-0,07
2 Vermittlung von mathematischem Grundwissen	0,04	0,85	0,05	0,05
3 Sichere Beherrschung von mathematischen Routineoperationen	0,02	0,73	0,02	-0,31
4 Automatisierte Beherrschung der wichtigsten Algorithmen der Mittelstufe	-0,14	0,32	0,03	-0,68
Schnelles Rechnen im Kopf bei Standardanwendungen	0,11	0,12	0,03	-0,61
6 Mathematisierungsfähigkeit entwickeln	-0,01	0,00	0,84	0,02
7 Alltagssituationen in mathematische Modelle zu übersetzen	0,06	-0,02	0,76	0,15
8 Mathematisch argumentieren können	0,18	0,24	0,60	0,01
9 Die Angemessenheit eines mathematischen Modells erkennen können	-0,13	-0,03	0,69	-0,31
10 Interesse für die Anwendungen der Mathematik	0,72	-0,18	0,01	-0,30
11 Mathematik im Alltag verstehen	0,62	-0,04	0,02	-0,45
12 Mathematisches Wissen im Alltag anwenden können	0,35	-0,02	0,21	-0,45
13 Freude am mathematischen Denken entwickeln	0,71	0,08	0,25	0,11
14 Interesse für die Anwendungen der Mathematik	0,79	-0,16	0,04	-0,17
15 Mathematik zu etwas persönlich Wichtigem machen	0,65	-0,01	0,18	0,24
16 Erreichen, dass die Schülerinnen und Schüler Spaß an der Mathematik haben	0,77	0,31	-0,15	0,14

Tabelle 29. Finale Dimensionen der Ziele im Mathematikunterricht für Gymnasien.

Item	Beherrschen von Routinen	Modellierungsfähigkeit	Anwendung im Alltag
1 Sicherheit bei der Bearbeitung von Standardaufgaben	0,86	0,01	0,13
2 Vermittlung von mathematischem Grundwissen	0,79	0,05	0,06
3 Sichere Beherrschung von mathematischen Routineoperationen	0,84	-0,01	-0,08
4 Automatisierte Beherrschung der wichtigsten Algorithmen der Mittelstufe	0,61	-0,02	-0,17
6 Mathematisierungsfähigkeit entwickeln	0,05	0,87	0,09
7 Alltagssituationen in mathematische Modelle zu übersetzen	0,16	0,75	-0,09
8 Mathematisch argumentieren können	-0,19	0,64	-0,09
9 Die Angemessenheit eines mathematischen Modells erkennen können	-0,11	0,68	0,00
11 Mathematik im Alltag verstehen	0,00	-0,01	-0,92
12 Mathematisches Wissen im Alltag anwenden können	-0,02	0,07	-0,87

12.4.2 Potential zur kognitive Aktivierung

12.4.2.1 Schulen mit mehreren Bildungsgängen

Tabelle 30. Faktorstruktur des Potentials zu kognitiven Aktivierung für Schulen mit mehreren Bildungsgängen.

	Kognitiv aktivierende Aufgaben	Insistieren auf Erklären und Begründen	Verständnis-fördernde Variation von Aufgaben-inhalten
1 Ich lasse Aufgaben bearbeiten, für die es keinen sofort erkennbaren Lösungsweg gibt.	-0,15	0,82	0,01
2 Ich stelle Fragen, die man nicht spontan beantworten kann, sondern die zum Nachdenken zwingen.	0,07	0,60	0,32
3 Ich lasse unterschiedliche Lösungswege von Aufgaben vergleichen und bewerten.	0,21	-0,21	0,59
4 Ich stelle Aufgaben, bei denen es nicht allein auf das Rechnen, sondern vor allem auf den richtigen Ansatz ankommt.	0,45	0,52	-0,02
5 Ich stelle Aufgaben, bei denen es auf das Verständnis des mathematischen Gedankengangs ankommt.	0,36	0,59	0,17
6 Ich stelle Aufgaben, die keine eindeutige Lösung haben, und lasse dies erklären.	-0,07	0,44	0,66
7 Ich stelle Aufgaben, für deren Lösung man Zeit zum Nachdenken braucht.	-0,17	0,25	0,79
8 Ich stelle auch Aufgaben, bei denen man mehrere Lösungswege zeigen muss.	0,68	0,36	-0,12
9 Ich halte Schüler/-innen dazu an, ihre Gedankengänge genau zu erklären.	0,81	-0,01	0,00
10 Ich verlange, dass Schüler/-innen ihre Arbeitsschritte ausführlich begründen.	0,66	-0,38	0,27
11 Bei mir wissen die Schüler/-innen, dass sie ihre Aussage immer begründen müssen, auch wenn ich es nicht extra dazusage.	0,89	-0,04	-0,09
12 In jeder Klassenarbeit bringe ich mindestens eine Aufgabe, in der eine Begründung gefordert ist.	0,47	0,46	0,17
13 Ich wechsle die Einkleidung von Aufgaben so, dass man erkennt, ob die mathematische Idee wirklich verstanden wurde.	0,78	0,04	0,07
14 Ich variiere Aufgaben so, dass man sieht, was die Schüler/-innen verstanden haben.	0,46	-0,20	0,42

Tabelle 31. Finale Dimensionen des Potentials zur kognitiven Aktivierung für Schulen mit mehreren Bildungsgängen.

	Kognitiv aktivierende Aufgaben	Insistieren auf Erklären und Begründen
9 Ich halte Schüler/-innen dazu an, ihre Gedankengänge genau zu erklären.	0,72	0,23
10 Ich verlange, dass Schüler/-innen ihre Arbeitsschritte ausführlich begründen.	0,88	-0,05
11 Bei mir wissen die Schüler/-innen, dass sie ihre Aussage immer begründen müssen, auch wenn ich es nicht extra dazusage.	0,93	-0,10
12 In jeder Klassenarbeit bringe ich mindestens eine Aufgabe, in der eine Begründung gefordert ist.	0,37	0,43
13 Ich wechsele die Einkleidung von Aufgaben so, dass man erkennt, ob die mathematische Idee wirklich verstanden wurde.	0,08	0,91
14 Ich variere Aufgaben so, dass man sieht, was die Schüler/-innen verstanden haben.	-0,12	0,96

12.4.2.2 Realschulen

Tabelle 32. Faktorstruktur des Potentials zu kognitiven Aktivierung für Realschulen.

	Kognitiv aktivierende Aufgaben	Insistieren auf Erklären und Begründen	Verständnis- fördernde Variation von Aufgaben- inhalten
1 Ich lasse Aufgaben bearbeiten, für die es keinen sofort erkennbaren Lösungsweg gibt.	0,76	0,10	-0,07
2 Ich stelle Fragen, die man nicht spontan beantworten kann, sondern die zum Nachdenken zwingen.	0,59	0,15	0,10
3 Ich lasse unterschiedliche Lösungswege von Aufgaben vergleichen und bewerten.	0,53	-0,02	0,09
4 Ich stelle Aufgaben, bei denen es nicht allein auf das Rechnen, sondern vor allem auf den richtigen Ansatz ankommt.	0,74	0,04	-0,10
5 Ich stelle Aufgaben, bei denen es auf das Verständnis des mathematischen Gedankengangs ankommt.	0,79	0,14	-0,10
6 Ich stelle Aufgaben, die keine eindeutige Lösung haben, und lasse dies erklären.	0,73	-0,31	0,15
7 Ich stelle Aufgaben, für deren Lösung man Zeit zum Nachdenken braucht.	0,73	0,04	-0,17
8 Ich stelle auch Aufgaben, bei denen man mehrere Lösungswege zeigen muss.	0,59	-0,12	0,36
9 Ich halte Schüler/-innen dazu an, ihre Gedankengänge genau zu erklären.	0,31	0,16	0,50
10 Ich verlange, dass Schüler/-innen ihre Arbeitsschritte ausführlich begründen.	-0,05	0,02	0,90
11 Bei mir wissen die Schüler/-innen, dass sie ihre Aussage immer begründen müssen, auch wenn ich es nicht extra dazusage.	-0,10	0,08	0,84
12 In jeder Klassenarbeit bringe ich mindestens eine Aufgabe, in der eine Begründung gefordert ist.	-0,03	0,60	0,11
13 Ich wechsle die Einkleidung von Aufgaben so, dass man erkennt, ob die mathematische Idee wirklich verstanden wurde.	0,15	0,82	0,00
14 Ich variiere Aufgaben so, dass man sieht, was die Schüler/-innen verstanden haben.	0,00	0,87	0,01

Tabelle 33. Finale Dimensionen des Potentials zur kognitiven Aktivierung für Realschulen.

	Kognitiv aktivierende Aufgaben	Insistieren auf Erklären und Begründen
9 Ich halte Schüler/-innen dazu an, ihre Gedankengänge genau zu erklären.	0,62	0,19
10 Ich verlange, dass Schüler/-innen ihre Arbeitsschritte ausführlich begründen.	0,93	-0,10
11 Bei mir wissen die Schüler/-innen, dass sie ihre Aussage immer begründen müssen, auch wenn ich es nicht extra dazusage.	0,82	-0,04
12 In jeder Klassenarbeit bringe ich mindestens eine Aufgabe, in der eine Begründung gefordert ist.	0,09	0,51
13 Ich wechsele die Einkleidung von Aufgaben so, dass man erkennt, ob die mathematische Idee wirklich verstanden wurde.	-0,06	0,91
14 Ich variiere Aufgaben so, dass man sieht, was die Schüler/-innen verstanden haben.	-0,05	0,90

12.4.2.3 Integrierte Gesamtschulen

Tabelle 34. Faktorstruktur des Potentials zu kognitiven Aktivierung für integrierte Gesamtschulen.

	Kognitiv aktivierende Aufgaben	Insistieren auf Erklären und Begründen	Verständnis- fördernde Variation von Aufgaben- inhalten
1 Ich lasse Aufgaben bearbeiten, für die es keinen sofort erkennbaren Lösungsweg gibt.	0,03	0,80	-0,03
2 Ich stelle Fragen, die man nicht spontan beantworten kann, sondern die zum Nachdenken zwingen.	0,70	0,28	-0,09
3 Ich lasse unterschiedliche Lösungswege von Aufgaben vergleichen und bewerten.	0,83	0,10	0,04
4 Ich stelle Aufgaben, bei denen es nicht allein auf das Rechnen, sondern vor allem auf den richtigen Ansatz ankommt.	0,13	0,73	0,30
5 Ich stelle Aufgaben, bei denen es auf das Verständnis des mathematischen Gedankengangs ankommt.	0,04	0,72	0,30
6 Ich stelle Aufgaben, die keine eindeutige Lösung haben, und lasse dies erklären.	0,45	0,67	0,02
7 Ich stelle Aufgaben, für deren Lösung man Zeit zum Nachdenken braucht.	0,66	0,39	-0,02
8 Ich stelle auch Aufgaben, bei denen man mehrere Lösungswege zeigen muss.	0,74	0,17	-0,16
9 Ich halte Schüler/-innen dazu an, ihre Gedankengänge genau zu erklären.	0,83	-0,24	0,00
10 Ich verlange, dass Schüler/-innen ihre Arbeitsschritte ausführlich begründen.	0,37	-0,32	0,19
11 Bei mir wissen die Schüler/-innen, dass sie ihre Aussage immer begründen müssen, auch wenn ich es nicht extra dazusage.	0,68	-0,05	0,19
12 In jeder Klassenarbeit bringe ich mindestens eine Aufgabe, in der eine Begründung gefordert ist.	-0,05	0,15	0,54
13 Ich wechsele die Einkleidung von Aufgaben so, dass man erkennt, ob die mathematische Idee wirklich verstanden wurde.	0,24	0,01	0,76
14 Ich variiere Aufgaben so, dass man sieht, was die Schüler/-innen verstanden haben.	-0,13	-0,01	0,91

Tabelle 35. Finale Dimensionen des Potentials zur kognitiven Aktivierung für integrierte Gesamtschulen.

	Kognitiv aktivierende Aufgaben	Insistieren auf Erklären und Begründen
9 Ich halte Schüler/-innen dazu an, ihre Gedankengänge genau zu erklären.	0,66	0,08
10 Ich verlange, dass Schüler/-innen ihre Arbeitsschritte ausführlich begründen.	0,81	-0,19
11 Bei mir wissen die Schüler/-innen, dass sie ihre Aussage immer begründen müssen, auch wenn ich es nicht extra dazusage.	0,75	0,19
12 In jeder Klassenarbeit bringe ich mindestens eine Aufgabe, in der eine Begründung gefordert ist.	-0,07	0,61
13 Ich wechsele die Einkleidung von Aufgaben so, dass man erkennt, ob die mathematische Idee wirklich verstanden wurde.	0,09	0,89
14 Ich variiere Aufgaben so, dass man sieht, was die Schüler/-innen verstanden haben.	0,06	0,79

12.4.2.4 Gymnasien

Tabelle 36. Faktorstruktur des Potentials zu kognitiven Aktivierung für Gymnasien.

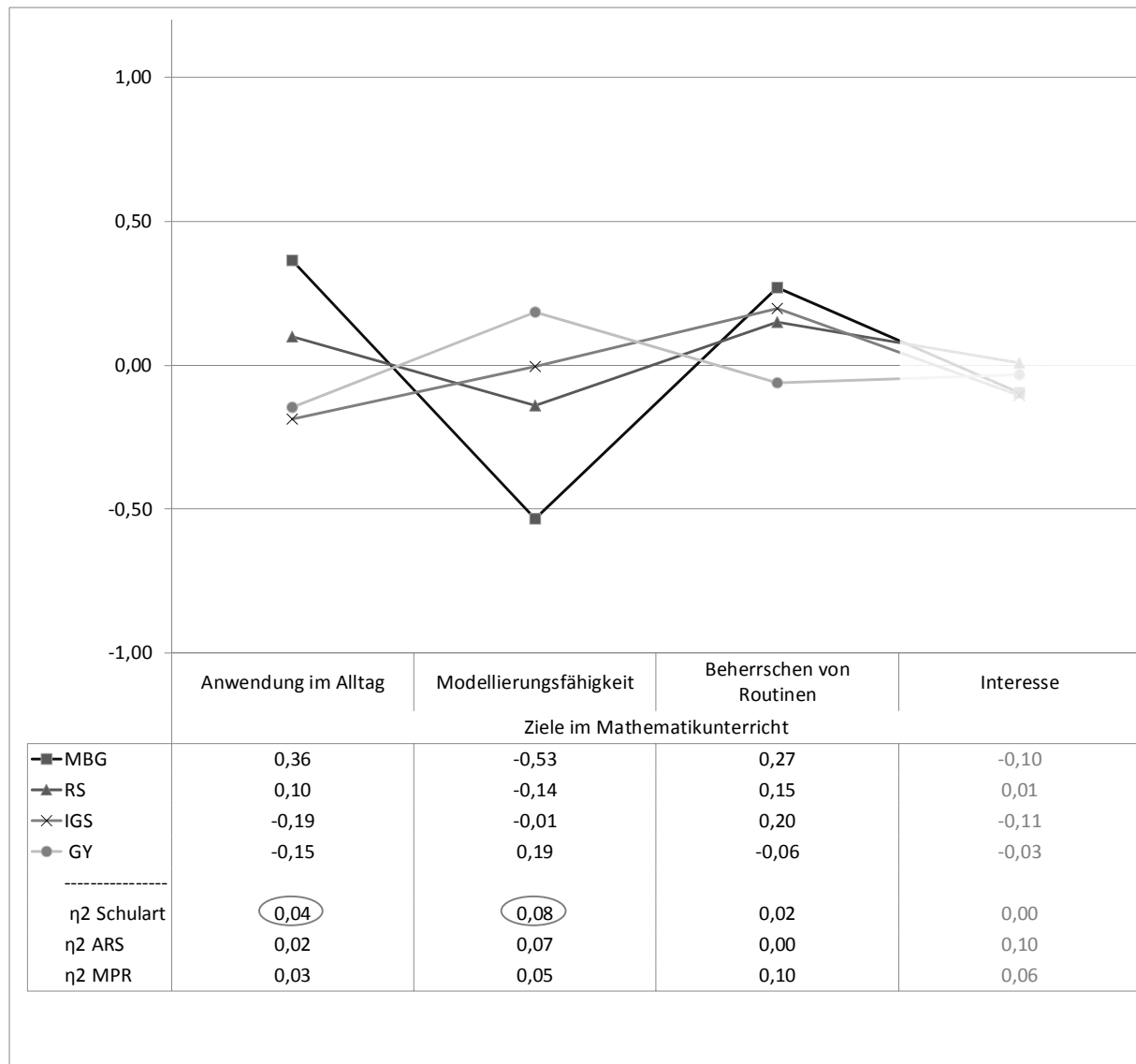
	Kognitiv aktivierende Aufgaben	Insistieren auf Erklären und Begründen	Verständnis- fördernde Variation von Aufgaben- inhalten
1 Ich lasse Aufgaben bearbeiten, für die es keinen sofort erkennbaren Lösungsweg gibt.	0,75	-0,16	-0,08
2 Ich stelle Fragen, die man nicht spontan beantworten kann, sondern die zum Nachdenken zwingen.	0,64	0,14	0,03
3 Ich lasse unterschiedliche Lösungswege von Aufgaben vergleichen und bewerten.	0,64	0,02	0,16
4 Ich stelle Aufgaben, bei denen es nicht allein auf das Rechnen, sondern vor allem auf den richtigen Ansatz ankommt.	0,21	0,15	-0,75
5 Ich stelle Aufgaben, bei denen es auf das Verständnis des mathematischen Gedankengangs ankommt.	0,26	0,16	-0,77
6 Ich stelle Aufgaben, die keine eindeutige Lösung haben, und lasse dies erklären.	0,68	-0,06	-0,21
7 Ich stelle Aufgaben, für deren Lösung man Zeit zum Nachdenken braucht.	0,63	0,16	-0,21
8 Ich stelle auch Aufgaben, bei denen man mehrere Lösungswege zeigen muss.	0,61	0,05	-0,03
9 Ich halte Schüler/-innen dazu an, ihre Gedankengänge genau zu erklären.	-0,02	0,75	-0,05
10 Ich verlange, dass Schüler/-innen ihre Arbeitsschritte ausführlich begründen.	0,04	0,78	0,07
11 Bei mir wissen die Schüler/-innen, dass sie ihre Aussage immer begründen müssen, auch wenn ich es nicht extra dazusage.	0,22	0,49	-0,08
12 In jeder Klassenarbeit bringe ich mindestens eine Aufgabe, in der eine Begründung gefordert ist.	-0,19	0,68	-0,11
13 Ich wechsle die Einkleidung von Aufgaben so, dass man erkennt, ob die mathematische Idee wirklich verstanden wurde.	0,14	0,72	0,12
14 Ich variiere Aufgaben so, dass man sieht, was die Schüler/-innen verstanden haben.	0,44	0,25	0,51

Tabelle 37. Finale Dimensionen des Potentials zur kognitiven Aktivierung für Gymnasien.

	Kognitiv aktivierende Aufgaben	Insistieren auf Erklären und Begründen
9 Ich halte Schüler/-innen dazu an, ihre Gedankengänge genau zu erklären.	0,84	-0,09
10 Ich verlange, dass Schüler/-innen ihre Arbeitsschritte ausführlich begründen.	0,73	0,20
11 Bei mir wissen die Schüler/-innen, dass sie ihre Aussage immer begründen müssen, auch wenn ich es nicht extra dazusage.	0,42	0,38
12 In jeder Klassenarbeit bringe ich mindestens eine Aufgabe, in der eine Begründung gefordert ist.	-0,15	0,70
13 Ich wechsele die Einkleidung von Aufgaben so, dass man erkennt, ob die mathematische Idee wirklich verstanden wurde.	0,50	0,51
14 Ich variiere Aufgaben so, dass man sieht, was die Schüler/-innen verstanden haben.	-0,14	0,95

12.5 Auswirkung von Antworttendenzen

12.5.1 Lehrerüberzeugungen



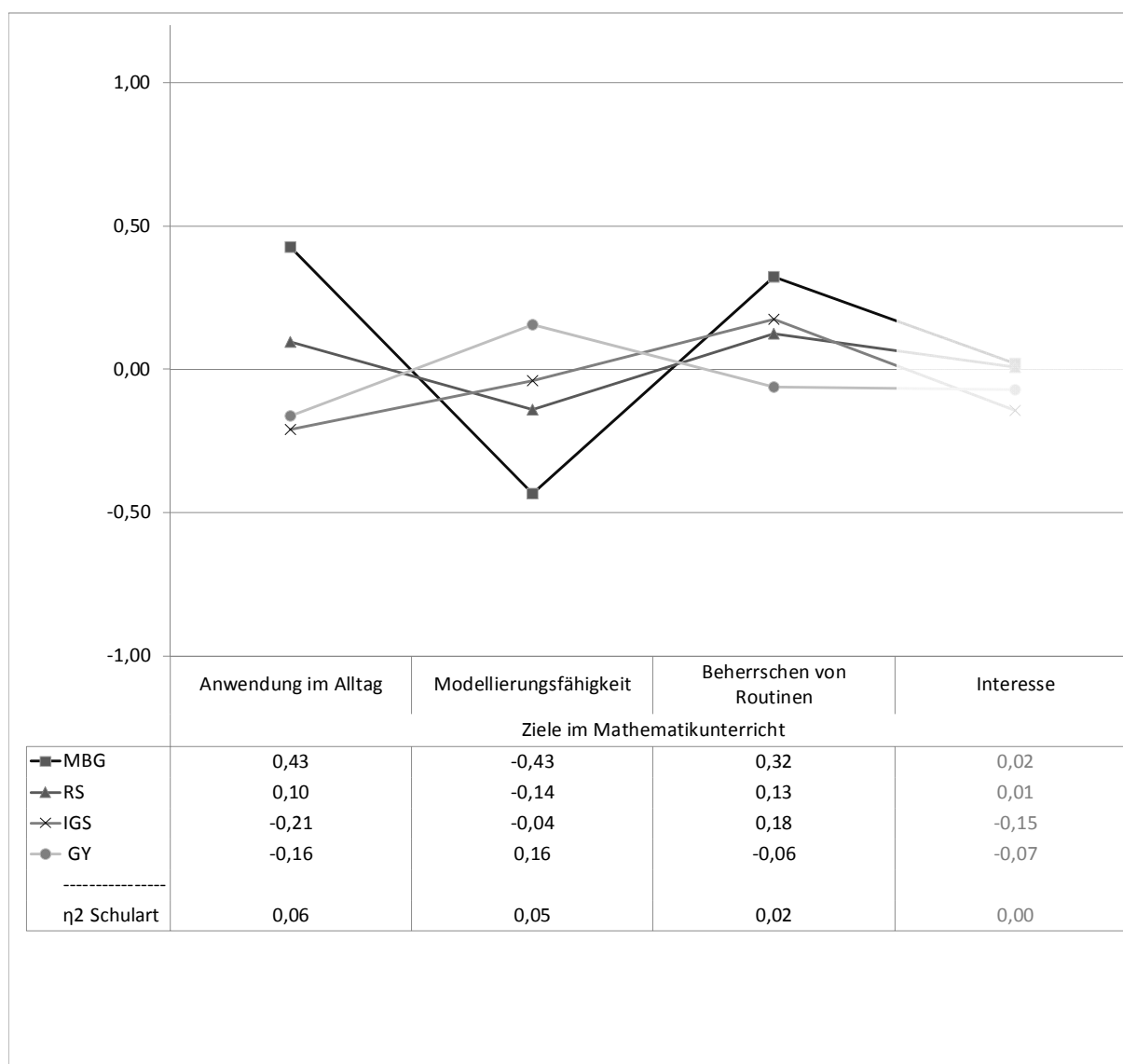
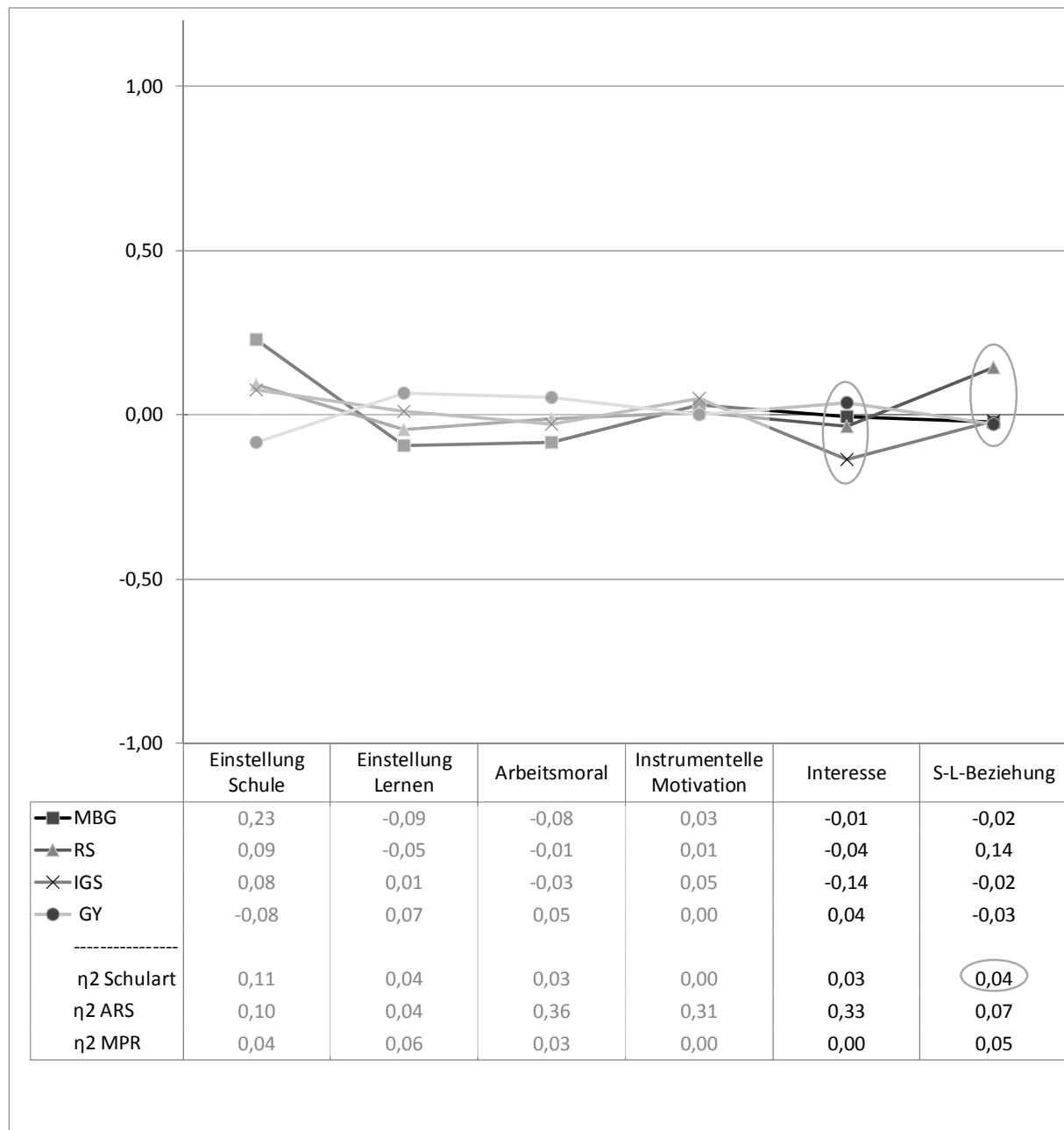


Abbildung 25. Lehrerüberzeugungen zu den Zielen des Mathematikunterrichts in den Schularten. Gegenüberstellung der Ergebnisse mit und ohne Berücksichtigung von Antworttendenzen.

Bedeutende Unterschiede sind eingekreist. Die Subdimension „Interesse“ ist nicht skalar invariant und daher grau unterlegt. Mittelwertvergleiche sind für diese Subdimension nicht zulässig.

Abkürzungen: MBG=Schulen mit mehreren Bildungsgängen, RS=Realschulen, IGS=integrierte Gesamtschulen, GY=Gymnasien, ARS=Akquieszenz, MPR=Mid-Point-Responding.

12.5.2 Affektive Merkmale



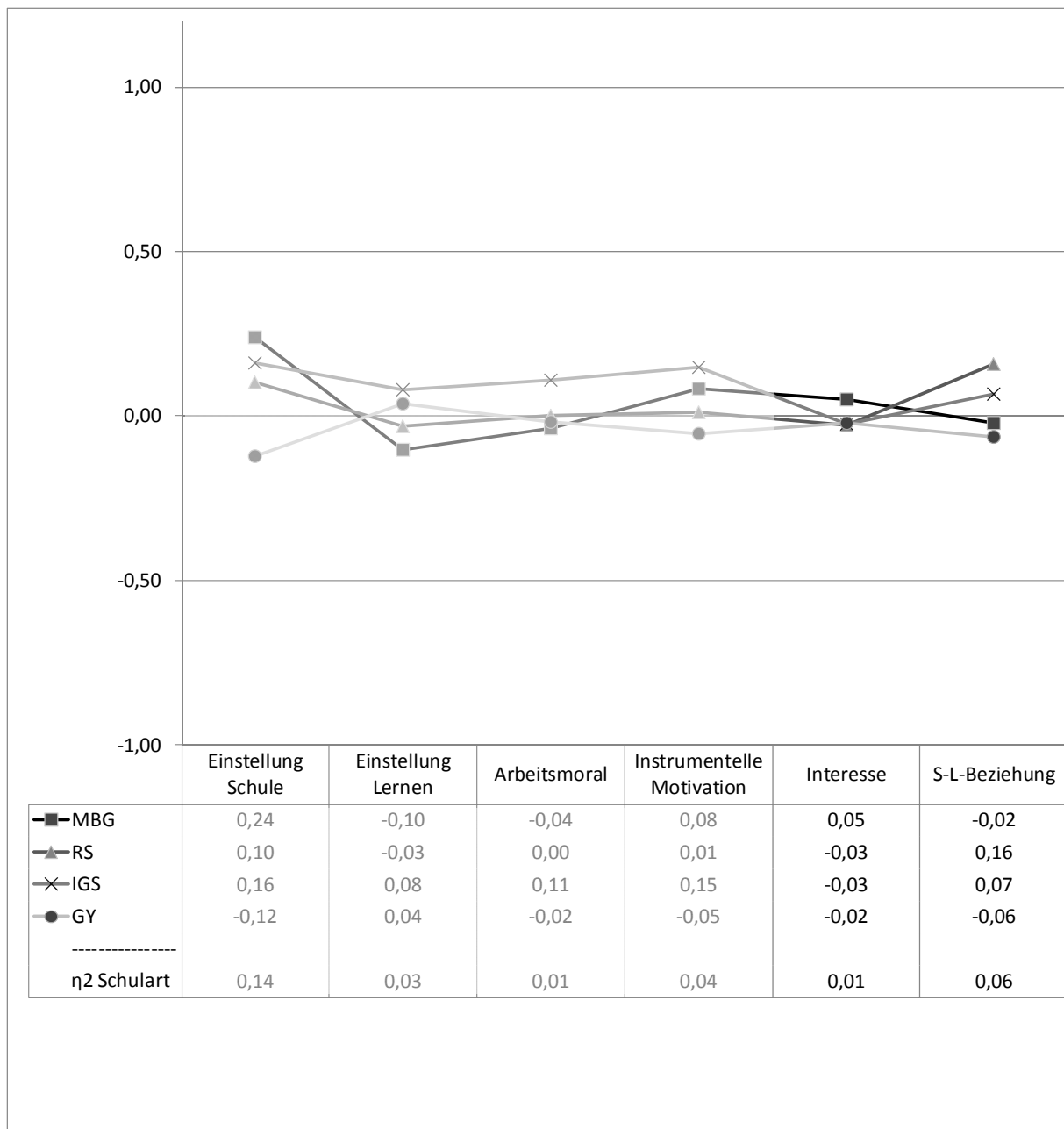
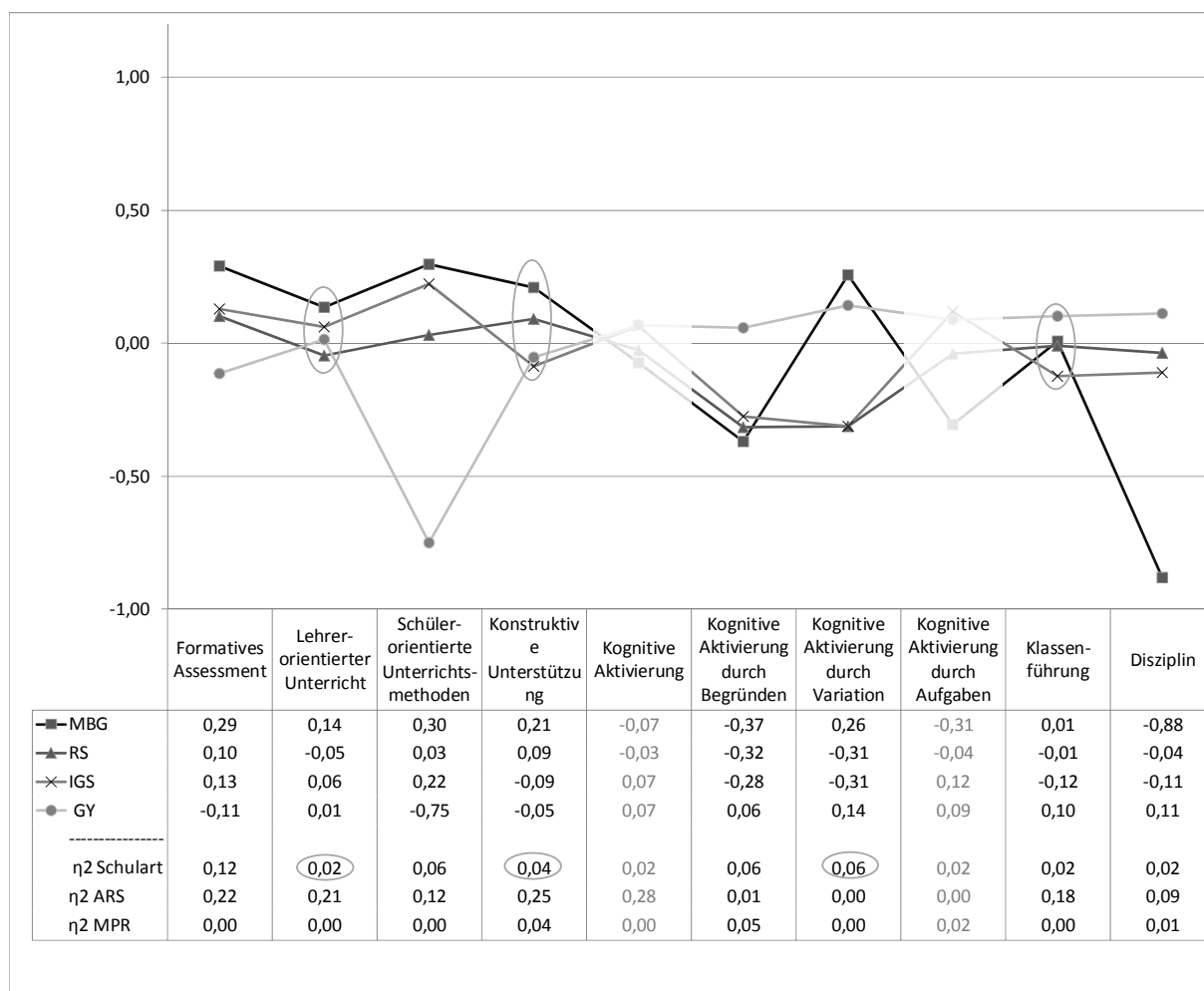


Abbildung 26. Mittlere Zusammensetzung der Klassen nach affektiven Merkmalen in den Schularten. Gegenüberstellung der Ergebnisse mit und ohne Berücksichtigung von Antworttendenzen.

Bedeutende Unterschiede sind eingekreist. Konstrukte, welche nicht skalar invariant sind, sind grau unterlegt. Mittelwertvergleiche sind für diese Konstrukte nicht zulässig.

Abkürzungen: MBG=Schulen mit mehreren Bildungsgängen, RS=Realschulen, IGS=integrierte Gesamtschulen, GY=Gymnasien, ARS=Akquieszenz, MPR=Mid-Point-Responding.

12.5.3 Unterrichtsprozesse



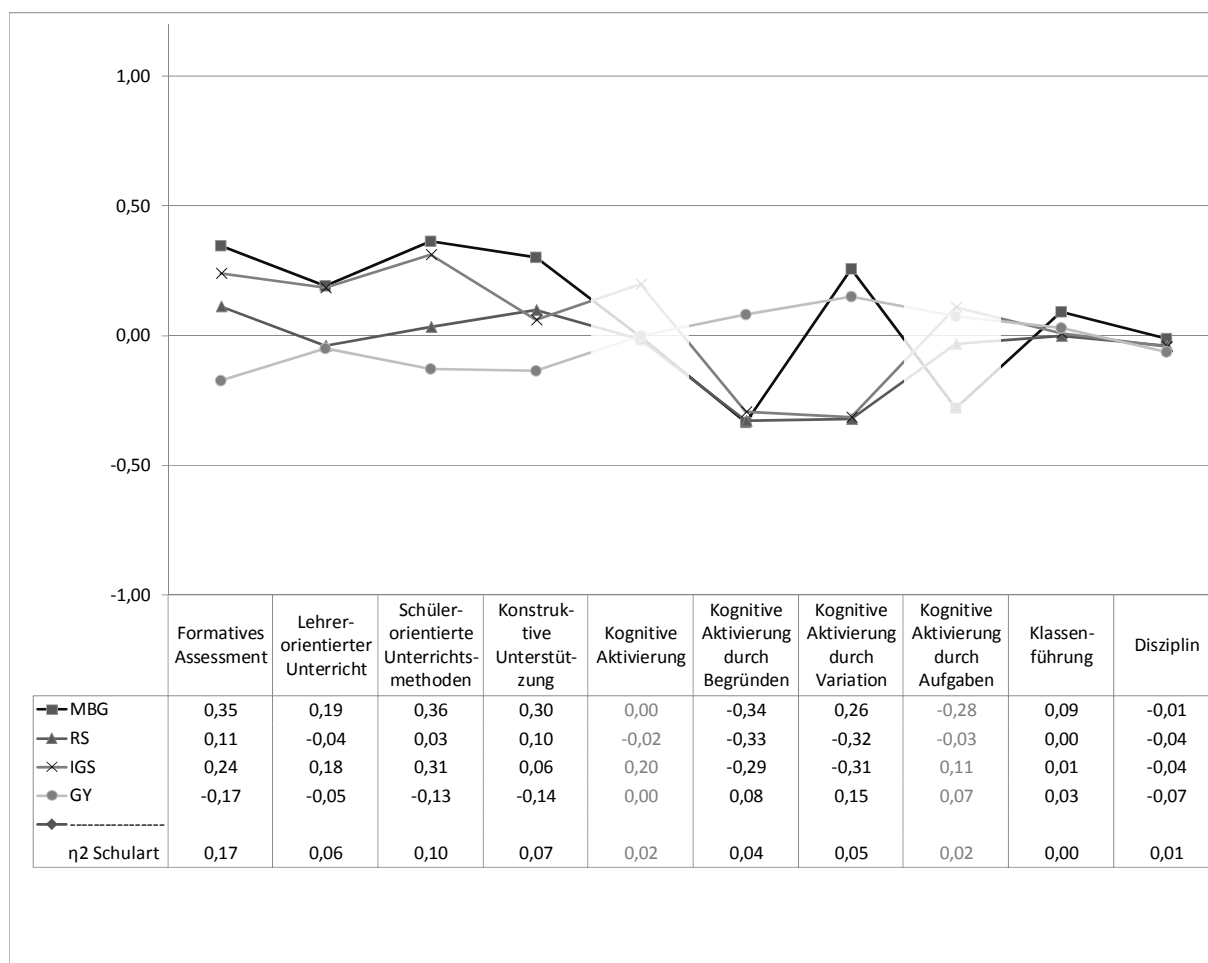


Abbildung 27. Unterrichtsprozesse im Fach Mathematik in den Schularten. Gegenüberstellung der Ergebnisse mit und ohne Berücksichtigung von Antworttendenzen.

Bedeutende Unterschiede sind eingekreist. Konstrukte, welche nicht skalar invariant sind, sind grau unterlegt. Mittelwertvergleiche sind für diese Konstrukte nicht zulässig.

Abkürzungen: MBG=Schulen mit mehreren Bildungsgängen, RS=Realschulen, IGS=integrierte Gesamtschulen, GY=Gymnasien, ARS=Akquieszenz, MPR=Mid-Point-Responding.

12.6 Mediationsmodelle

12.6.1 Mediation der Schulartunterschiede im formativen Assessment

Tabelle 38. Mediation der Schulartunterschiede im formativen Assessment

		Unterrichtsziel: Modellierungsfähigkeit	Mathematische Konzepte	Reine Mathematik	Leistungskomposition	Soziale Komposition
		Konfidenz- β intervall	Konfidenz- β intervall	Konfidenz- β intervall	Konfidenz- β intervall	Konfidenz- β intervall
Intercept	Formatives Assessment	-0,11 *	-0,10 *	-0,12 *	-0,07	-0,07
	Mediator	0,15	0,35 *	0,21 *	0,58 *	0,33 *
Slope	Direkt	RS	0,24 *	0,16 *	0,20 *	0,10
		IGS	0,24	0,17	0,21	0,11
	Formatives Assessment	MBG	0,38 *	0,33 *	0,38 *	0,26
		Mediator	0,03	-0,10	-0,05	-0,10
	Direkt	RS	-0,31 *	-0,63 *	-0,34 *	-1,12 *
		IGS	-0,20 *	-0,80 *	-0,52 *	-1,32 *
	Mediator	MBG	-0,62 *	-0,77 *	-0,46 *	-1,44 *
						0,80 *
	Indirekt	RS	-0,01 [-0,05;0,00]	0,06 [-0,02;0,14]	0,02 [-0,03;0,06]	0,11 [-0,07;0,30]
		IGS	-0,01 [-0,05;0,01]	0,08 [-0,02;0,16]	0,03 [-0,06;0,10]	0,13 [-0,08;0,34]
		MBG	-0,02 [-0,08;0,01]	0,08 [-0,02;0,17]	0,02 [-0,04;0,08]	0,15 [-0,09;0,37]
R ² Klasse	Mediator	0,05	0,52 *	0,27 *	0,37 *	0,56 *
R ² Klasse	Formatives Assessment	0,37 *	0,37 *	0,23 *	0,81 *	0,37 *

Signifikante direkte Effekte ($p < 0,05$) sind mit einem Stern gekennzeichnet (*). Für die indirekten Effekte sind die 95 % Bootstrap Konfidenzintervalle angegeben. Enthält das Bootstrap-Vertrauensintervall den Wert 0 nicht, sind signifikante indirekten Effekt anzunehmen.

Abkürzungen: MBG=Schulen mit mehreren Bildungsgängen, RS=Realschulen, IGS=integrierte Gesamtschulen, Direkt=direkter Effekt, Indirekt=indirekter Effekt

12.6.2 Mediation der Schulartunterschiede in schülerorientierten Unterrichtsmethoden

Tabelle 39. Mediation der Schulartunterschiede in schülerorientierten Unterrichtsmethoden.

			Unterrichtsziel: Modellierungsfähigkeit	Mathematische Konzepte	Reine Mathematik	Leistungskomposition	Soziale Komposition
			Konfidenz- intervall β	Konfidenz- intervall β	Konfidenz- intervall β	Konfidenz- intervall β	Konfidenz- intervall β
Intercept	Schülerorient. Methoden		0,24 *	0,00	-0,03	0,13	-0,06
	Mediator		1,03 *	0,35 *	0,21 *	0,58 *	0,33 *
Slope	Direkt Schülerorientierung	RS	0,14	-0,05	0,01	-0,32	0,04
		IGS	0,29	0,11	0,14	-0,20	0,23
		MBG	0,35 *	0,18	0,23	-0,18	0,30
		Mediator	-0,03	-0,25 *	-0,27 *	-0,38 *	-0,10
	Direkt Mediator	RS	-0,31 *	-0,63 *	-0,34 *	-1,12 *	-0,68 *
		IGS	-0,20 *	-0,80 *	-0,51 *	-1,32 *	-0,70 *
		MBG	-0,61 *	-0,77 *	-0,46 *	-1,44 *	-0,80 *
	Indirekt	RS	0,01 [-0,01;0,06]	0,16 * [0,02;0,33]	0,09 * [0,03;0,17]	0,43 * [0,19;0,70]	0,07 [-0,10;0,23]
		IGS	0,01 [-0,01;0,06]	0,20 * [0,03;0,41]	0,14 * [0,05;0,28]	0,50 * [0,23;0,81]	0,07 [-0,10;0,24]
		MBG	0,02 [-0,03;0,10]	0,12 * [0,04;0,40]	0,13 * [0,04;0,13]	0,55 * [0,26;0,89]	0,08 [-0,12;0,27]
R ² Klasse	Mediator		0,05	0,22 *	0,26	0,81 *	0,21 *
R ² Klasse	Schülerorient. Methoden		0,27	0,52 *	0,26	0,24 *	0,56 *

Signifikante direkte Effekte ($p < 0,05$) sind mit einem Stern gekennzeichnet (*). Für die indirekten Effekte sind die 95 % Bootstrap Konfidenzintervalle angegeben. Enthält das Bootstrap-Vertrauensintervall den Wert 0 nicht, sind signifikante indirekten Effekt anzunehmen.

Abkürzungen: MBG=Schulen mit mehreren Bildungsgängen, RS=Realschulen, IGS=integrierte Gesamtschulen, Direkt=direkter Effekt, Indirekt=indirekter Effekt

12.6.3 Mediation der Schulartunterschiede im Potential zur kognitiven Aktivierung durch Insistieren auf Begründen

Tabelle 40. Mediation der Schulartunterschiede im Potential zur kognitiven Aktivierung durch Insistieren auf Begründen.

			Unterrichtsziel: Modellierungsfähigkeit	Mathematische Konzepte	Reine Mathematik	Leistungskomposition	Soziale Komposition
			Konfidenz- intervall	Konfidenz- intervall	Konfidenz- intervall	Konfidenz- intervall	Konfidenz- intervall
			β	β	β	β	β
Intercept	Kog. Akt. Begründen		0,09	0,13	0,24 *	0,32	0,25
	Mediator		0,15	0,33 *	0,25 *	0,60 *	0,33 *
Slope	Direkt	RS	-0,31	-0,37	-0,55 *	-0,72	-0,62 *
		IGS	-0,31	-0,32	-0,55	-0,72	-0,58
		MBG	-0,28	-0,41	-0,65 *	-0,88	-0,70
		Mediator	0,31 *	0,08	-0,36	-0,27	-0,30
	Direkt	RS	-0,31 *	-0,61 *	-0,36 *	-1,11 *	-0,67 *
		IGS	-0,20 *	-0,04 *	-0,49 *	-1,26 *	-0,68 *
		MBG	-0,61 *	-0,74 *	-0,49 *	-1,47 *	-0,77 *
	Indirekt	RS	-0,10 [-0,25;0,00]	-0,05 [-0,29;0,22]	0,13 [-0,02;0,29]	0,30 [-0,21;0,81]	0,20 [-0,14;0,53]
		IGS	-0,06 [-0,23;0,07]	-0,06 [-0,34;0,25]	0,18 [-0,02;0,45]	0,34 [-0,24;0,93]	0,20 [-0,14;0,58]
		MBG	-0,19 [-0,46;0,04]	-0,06 [-0,36;0,26]	0,14 [-0,02;0,38]	0,40 [-0,26;1,09]	0,23 [-0,15;0,64]
	Mediator		0,05	0,49 *	0,25 *	0,14	0,52 *
	Kog. Akt. Begründen		0,18 *	0,13	0,14 *	0,81 *	0,13

Signifikante direkte Effekte ($p < 0,05$) sind mit einem Stern gekennzeichnet (*). Für die indirekten Effekte sind die 95 % Bootstrap Konfidenzintervalle angegeben. Enthält das Bootstrap-Vertrauensintervall den Wert 0 nicht, sind signifikante indirekten Effekt anzunehmen.

Abkürzungen: MBG=Schulen mit mehreren Bildungsgängen, RS=Realschulen, IGS=integrierte Gesamtschulen, Direkt=direkter Effekt, Indirekt=indirekter Effekt

12.6.4 Mediation der Schulartunterschiede im Potential zur kognitiven Aktivierung durch Variation der Aufgaben

Tabelle 41. Mediation der Schulartunterschiede im Potential zur kognitiven Aktivierung durch Variation der Aufgaben.

			Unterrichtsziel: Modellierungsfähigkeit	Mathematische Konzepte	Reine Mathematik	Leistungskomposition	Soziale Komposition
			Konfidenz- β intervall	Konfidenz- β intervall	Konfidenz- β intervall	Konfidenz- β intervall	Konfidenz- β intervall
Intercept	Kog. Akt. Variation		0,10	-0,06	0,11	-0,02	-0,02
	Mediator		0,15	0,33 *	0,25 *	0,60 *	0,33 *
Slope	Direkt Kog. Akt. Variation	RS	-0,38	-0,07	-0,42 *	-0,16	-0,11
		IGS	-0,43 *	0,00	-0,39	-0,12	-0,10
		MBG	0,28	0,67	0,22	0,57	0,56
		Mediator	0,31	0,69 *	0,19	0,29	0,56
	Direkt Mediator	RS	-0,31 *	-0,61 *	-0,36 *	-1,11 *	-0,67 *
		IGS	-0,20 *	-0,70 *	-0,48 *	-1,26 *	-0,68 *
		MBG	-0,48 *	-0,74 *	-0,49 *	-1,47 *	-0,77 *
	Indirekt	RS	-0,10 [-0,41;-0,01]	-0,42 * [-0,79;-0,22]	-0,07 [-0,28;0,05]	-0,33 [-0,98;0,17]	-0,38 [-1,16;0,19]
		IGS	-0,06 [-0,41;0,01]	-0,49 * [-0,93;-0,23]	-0,09 [-0,44;0,07]	-0,37 [-1,11;0,20]	-0,38 [-1,27;0,17]
		MBG	-0,15 [-0,50;0,03]	-0,52 * [-1,05;-0,21]	-0,09 [-0,41;0,06]	-0,43 [-1,35;0,29]	-0,43 [-1,24;0,22]
R ² Klasse	Mediator		0,03	0,49	0,25 *	0,06	0,52 *
R ² Klasse	Kog. Akt. Variation		0,13	0,10 *	0,06	0,81 *	0,08

Signifikante direkte Effekte ($p < 0,05$) sind mit einem Stern gekennzeichnet (*). Für die indirekten Effekte sind die 95 % Bootstrap Konfidenzintervalle angegeben. Enthält das Bootstrap-Vertrauensintervall den Wert 0 nicht, sind signifikante indirekten Effekt anzunehmen

Abkürzungen: MBG=Schulen mit mehreren Bildungsgängen, RS=Realschulen, IGS=integrierte Gesamtschulen, Direkt=direkter Effekt, Indirekt=indirekter Effekt.